

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局



(43) 国際公開日  
2005年7月7日 (07.07.2005)

PCT

(10) 国際公開番号  
WO 2005/062293 A1

(51) 国際特許分類<sup>7</sup>: G10L 15/00, 15/22, 15/28, G11B 20/10

灘区徳井町四丁目2番21号徳井ハイツ205号  
Hyogo (JP).

(21) 国際出願番号: PCT/IB2004/004007

(74) 代理人: 木村 满 (KIMURA,Mitsuru); 〒101-0054 東京都千代田区神田錦町二丁目7番地 協販ビル2階  
Tokyo (JP).

(22) 国際出願日: 2004年12月6日 (06.12.2004)

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) 国際出願の言語: 日本語

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ヨーラシア (AM, AZ,

(26) 国際公開の言語: 日本語

/ 続葉有 /

(30) 優先権データ:  
特願2003-406669 2003年12月5日 (05.12.2003) JP

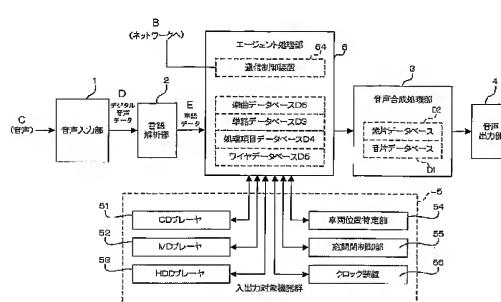
(71) 出願人(米国を除く全ての指定国について): 株式会社ケンウッド (KABUSHIKI KAISHA KENWOOD) [JP/JP]; 〒192-8525 東京都八王子市石川町2967-3 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 小山 利佳 (KOYAMA,Rika) [JP/JP]; 〒657-0033 兵庫県神戸市

(54) Title: AUDIO DEVICE CONTROL DEVICE, AUDIO DEVICE CONTROL METHOD, AND PROGRAM

(54) 発明の名称: オーディオ機器制御装置、オーディオ機器制御方法及びプログラム



B... (TO NETWORK)  
C... (SPEECH)  
1... SPEECH INPUT UNIT  
D... DIGITAL SPEECH DATA  
2... LANGUAGE ANALYSIS UNIT  
E... WORD DATA  
6... AGENT PROCESSING UNIT  
64... COMMUNICATION CONTROL DEVICE  
D8... MUSIC COMPOSITION DATABASE  
D3... WORD DATABASE  
D4... PROCESSING ITEM DATABASE  
D5... WIRE DATABASE

3... SPEECH SYNTHESIS PROCESSING UNIT  
D2... PHONEME ELEMENT DATABASE  
D1... PHONEME ELEMENT DATABASE  
4... SPEECH OUTPUT UNIT  
51... CD PLAYER  
52... MD PLAYER  
53... HDD PLAYER  
5... I/O OBJECT DEVICE GROUP  
54... VEHICLE POSITION SPECIFICATION UNIT  
55... WINDOW OPEN/CLOSE CONTROL UNIT  
56... CLOCK DEVICE

(57) Abstract: A language analysis unit (2) subjects the speech inputted by a speech input unit (1) to speech recognition so as to specify a word and its score which may be expressed by the speech and supplies the word data expressing them to an agent processing unit (6). The agent processing unit (6) contains: processing item data defining a data acquiring process for acquiring word data and the like, judgment process, and I/O process; and a wire which is data defining the transition from a process to another and giving a weight coefficient to the transition. By executing the flow expressed by the processing item data and the wire as a whole, devices belonging to the I/O object device group (5) are controlled. To which process in the flow the transition proceeds is decided by the weight coefficient of each wire decided by the point where the process has advanced and the wire connection relationship and the score of the word data. The wire and the processing item data can be downloaded from an external server.

/ 続葉有 /

WO 2005/062293 A1



BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

- **国際調査報告書**
- 請求の範囲の補正の期限前の公開であり、補正書受領の際には再公開される。

---

(57) 要約:

言語解析部2は、音声入力部1が入力した音声に音声認識を施して、この音声が表している可能性のある単語とそのスコアとを特定し、これらを表す単語データをエージェント処理部6に供給する。エージェント処理部6は、単語データ等を取得するデータ取得処理や、判別処理、入出力処理を定義する処理項目データと、ひとつの処理から次の処理への遷移を定義しこの遷移に重み係数を与えるデータであるワイヤとを記憶しており、処理項目データとワイヤとが全体として表すフローを実行することにより、入出力対象機器群5に属する機器を制御する。フロー内のどの処理へと遷移するかは、処理が進んだ地点とワイヤの接続関係とにより決まる各ワイヤの重み係数と、単語データのスコアとにより決まる。ワイヤや処理項目データは外部サーバよりダウンロード可能である。

## 明細書

### オーディオ機器制御装置、オーディオ機器制御方法及びプログラム技術分野

[0001] この発明は、オーディオ機器制御装置、オーディオ機器制御方法、データ出力装置及びプログラムに関する。

#### 背景技術

[0002] 近年、音声認識の技術を用いて音声を認識し、認識結果に応答して電気機器などを制御する手法が用いられている。この手法は、具体的には、入力した音声が表す単語を識別し、識別された単語が所定のキーワードに合致するか否かを判別して、判別結果に基づいて外部の機器を制御するものである。このような手法は、例えばオーディオ機器の制御にも用いられている(例えば、特許文献1及び特許文献2参照)。

特許文献1:特開平4-324312号公報

特許文献2:特開2002-182688号公報

#### 発明の開示

##### 発明が解決しようとする課題

[0003] しかし、人間が言語の形で発する指示を完全に認識するのは困難である。このため、上述の手法では、人間が言語の形で発する指示に適切に応答することができない場合があった。

[0004] この発明は上記実状に鑑みてなされたものであり、人間が言語の形で発する指示に適切に応答して機器を制御できるオーディオ機器制御装置、オーディオ機器制御方法及びプログラムを提供することを目的とする。

#### 課題を解決するための手段

[0005] 上記目的を達成するため、この発明の第1の観点にかかるオーディオ機器制御装置は、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、

制御対象であるオーディオ機器の動作状態を示す動作状態情報を取得する動作状態情報取得手段と、

前記音声認識手段が特定した候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、前記動作状態情報取得手段が取得した動作状態情報に基づいて、当該利用態様を実現するために当該オーディオ機器に加えるべき操作を特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、

ことを特徴とする。

[0006] また、この発明の第2の観点にかかるオーディオ機器制御装置は、音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、

前記音声認識手段が特定した候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、

ことを特徴とする。

[0007] また、この発明の第3の観点にかかるオーディオ機器制御装置は、音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、

前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定手段と、

前記音声認識手段が特定した候補と、前記発話者特定手段が特定した発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を

加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、  
ことを特徴とする。

[0008] 前記環境データは、例えば、前記オーディオ機器が使用される環境の現在位置を示すデータからなっている。

[0009] 前記環境データは、例えば、前記オーディオ機器が使用される環境の遮音状況を示すデータからなっている。

[0010] また、この発明の第4の観点にかかるオーディオ機器制御方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
制御対象であるオーディオ機器の動作状態を取得する動作状態取得ステップと、  
前記音声認識ステップで特定された候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、動作状態取得ステップで取得し動作状態とにに基づき、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、  
ことを特徴とする。

[0011] また、この発明の第5の観点にかかるオーディオ機器制御方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
前記音声認識ステップで特定された候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとにに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、  
ことを特徴とする。

[0012] また、この発明の第6の観点にかかるオーディオ機器制御方法は、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当

該音声が表す語句の候補を特定する音声認識ステップと、

前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定ステップと、

前記音声認識ステップで特定された候補と、前記発話者特定ステップで特定された発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、

ことを特徴とする。

[0013] また、この発明の第7の観点に係るコンピュータプログラムは、  
コンピュータに、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
制御対象であるオーディオ機器の動作状態を取得する動作状態取得ステップと、  
前記音声認識ステップで特定された候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、動作状態取得ステップで取得し動作状態とにに基づき、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、を実行させる ことを特徴とする。

[0014] また、この発明の第8の観点に係るコンピュータプログラムは、  
コンピュータに、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
前記音声認識ステップで特定された候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとにに基づき、前記音声の発話者が欲する当該オーディオ機器の利用

態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、

を実行させることを特徴とする。

[0015] また、この発明の第9の観点に係るコンピュータプログラムは、コンピュータに、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、

前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定ステップと、

前記音声認識ステップ特定された候補と、前記発話者特定ステップで特定された発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、

を実行させる。

### 発明の効果

[0016] この発明によれば、人間が言語の形で発する指示に適切に応答して機器を制御できるオーディオ機器制御装置、オーディオ機器制御方法及びプログラムが実現される。

### 図面の簡単な説明

[0017] [図1]この発明の実施の形態に係るオーディオ機器制御システムを示す図である。

[図2]この実施の形態に係るオーディオ機器制御システムのより詳細な構成図である。

。

[図3]楽曲データベースのデータ構成の例を示す図である。

[図4]単語データベースの具体例を模式的に示す図である。

[図5]トリガ取得処理を説明するための図である。

[図6]判別処理を説明するための図である。

[図7]問い合わせ付きの判別処理を説明するための図である。

[図8]入出力処理を説明するための図である。

[図9]ワイヤの例を示す図である。

[図10]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図11]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図12]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図13]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図14]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図15]処理項目データベース及びワイヤデータベースが全体として表しているフローを示す図である。

[図16]各ワイヤに設定される重み係数Jを説明するための図である。

[図17]エージェント処理部の動作を説明するためのフローチャートである。

[図18]外部サーバ(配信サーバ)の構成と、ネットワーク構成とを示す図である。

[図19]ダウンロード処理を説明するためのフローチャートである。

### 符号の説明

- [0018] 1 音声入力部
- 2 言語解析部
- 3 音声合成処理部
- 4 音声出力部
- 5 入出力対象機器群
- 51 CDプレーヤ
- 52 MDプレーヤ

- 53 HDDプレーヤ
- 54 車両位置特定部
- 55 窓開閉制御部
- 56 クロック
- 6 エージェント処理部

### 発明を実施するための最良の形態

[0019] 以下、図面を参照して、この発明の実施の形態を、車両内に設置されたオーディオ機器制御システムを例として説明する。

図1は、このオーディオ機器制御システムの構成を示すブロック図である。また、図2は、各部の物理的な構成の例を示すブロック図である。

図1に示すように、このオーディオ機器制御システムは、音声入力部1と、言語解析部2と、音声合成処理部3と、音声出力部4と、入出力対象機器群5と、エージェント処理部6より構成されている。

[0020] 音声入力部1は、音声を入力し、入力した音声からデジタル形式の音声データを生成し、この音声データを言語解析部2へと供給する。具体的には、音声入力部1は、図2に示すように、例えば、マイクロフォン11と、AF (Audio Frequency) 増幅器12と、サンプルホールド回路を内蔵するA／D (Analog-to-Digital) コンバータ13などより構成されている。マイクロフォン11は音声を音声信号に変換して出力する。AF増幅器12は、マイクロフォン11からの音声信号を増幅して出力する。A／Dコンバータ13は、AF増幅器12からの増幅された音声信号をサンプリング、A／D変換することにより、デジタル音声データを生成し、言語解析部2へと供給する。

[0021] 言語解析部2と音声合成処理部3とエージェント処理部6とは、図2に示すように、それぞれ、例えば、CPU (Central Processing Unit) 等からなるプロセッサ21、31、61と、このプロセッサ21、31、61が実行するプログラムを記憶するハードディスク装置等の不揮発性メモリ22、32、62と、プロセッサのワークエリアとなる記憶領域を有するRAM (Random Access Memory) 等の揮発性メモリ23、33、63より構成されている。なお、言語解析部2、音声合成処理部3及びエージェント処理部6の一部又は全部の機能を1つのプロセッサや1つの不揮発性メモリや1つの揮発性メモリで構成しても

よい。

[0022] 言語解析部2は、音声入力部1より供給された音声データに音声認識処理を行う。言語解析部2は、音声認識処理により、音声データが表している単語の候補と、この候補の尤度(スコア)とを特定する。音声認識の手法は任意である。また、単語の候補は複数特定されてよい。言語解析部2は、特定した候補とその候補のスコアとを示すデータ(以下、単語データと呼ぶ)を生成し、エージェント処理部6へと供給する。

[0023] 音声合成処理部3の不揮発性メモリ32は、単語の波形を表すデータを記憶する音片データベースD1と、音素を構成するための波形を表すデータを記憶する素片データベースD2とを記憶する。

[0024] 音片データベースD1は、単語の波形を表すデータを記憶する。素片データベースD2は、音素を構成するための波形を表すデータを記憶する。音声合成処理部3は、音片データベースD1及び／又は素片データベースD2に格納されているデータを用いて、エージェント処理部6より供給された文章データを読み上げる音声を表すデジタル音声データを生成する。

[0025] 音声合成部3は、生成した音声データを音声出力部4に供給する。デジタル音声データを生成する手法は任意であるが、例えば、録音編集方式や規則合成方式(Rule-based synthesis)を使用できる。なお、録音編集方式は、例えば、単語単位の音声を予めアナウンサーに読んでもらい、それらをつなぎあわせて出力する方式である。また、規則合成方式は、音韻(子音や母音)や仮名のような比較的小な単位をつなぎあわせて出力する方式である。

[0026] 音声出力部4は、音声合成処理部3から供給されたデジタル音声データが表す音声を再生する。より詳細には、音声出力部4は、図2に示すように、D/A(Digital-to-Analog)コンバータ41、AF増幅器42及びスピーカ43を備える。D/Aコンバータ41は、音声合成処理部3より供給されたデジタル音声データをD/A変換して、アナログ音声信号に変換する。AF増幅器42は、アナログ音声信号を増幅する。スピーカ43は、アナログ音声信号に従って振動し、アナログ音声データが表す音声を再生し、放音する。

[0027] 入出力対象機器群5は、例えば、CD(Compact Disc)プレーヤ51、MD(Mini Disc

)プレーヤ52、HDD(ハードディスクドライブ)プレーヤ53、車両位置特定部54、窓開閉制御部55や、クロック装置56、などを含む。

[0028] CDプレーヤ51は、自己に装着されたCDに記録された音声データが表す音声の再生や再生の停止、ミュート、早送り再生、巻き戻し再生、一時停止、CDの排出、等の動作を、エージェント処理部6が供給する制御信号に従って行う。

MDプレーヤ52は、自己に装着されたMDに記録された音声データが表す音声の再生や再生の停止、ミュート、早送り再生、巻き戻し再生、一時停止、CDの排出、等の動作を、エージェント処理部6が供給する制御信号に従って行う。

HDDプレーヤ53は、不揮発性の磁気ディスクであり、音声データの記憶、音声データが表す音声の再生や再生の停止、ミュート、早送り再生、巻き戻し再生、一次停止、等の動作を、エージェント処理部6が供給する制御信号に従って行う。

[0029] 車両位置特定部54は、GPS(Global Positioning System)の移動局を用いた公知のカーナビゲーションシステムなどより構成されている。車両位置特定部54は、例えば、エージェント処理部6が供給する制御信号に従って車両の現在位置を検知し、検知した現在位置が地図上でどの位置に当たるかを表すデータを生成して、エージェント処理部6に供給する。

[0030] 窓開閉制御部55は、モータや、制御信号に従ってモータの回転及び停止を制御する制御回路や、モータの回転に従って車両の窓ガラスを動かすためのウインチ等から構成されている。窓開閉制御部55は、自己に供給される制御信号に従って、車両の窓の開閉を行う。また、窓開閉制御部55の制御回路は、窓開閉制御部55の動作状態を表すデータとして、例えば、開閉する対象の窓が開いている量を示すデータを生成して出力する。

[0031] クロック装置56は、例えば水晶発振器やカウンタ回路等より構成されておいる。クロック装置56は、現在の時刻を示すデータを連続的に生成し、エージェント処理部6に供給する。尚、クロック装置56は、カーナビゲーションシステムを介してGPSから得られる時刻情報を用いてもよい。

[0032] エージェント処理部6は、モ뎀やパケット通信端末等からなる通信制御装置64を備え、この通信制御装置を介して外部のネットワーク(例えば、無線電話回線を介し

たインターネット)を介して後述する配信サーバ100に接続している。

[0033] また、エージェント処理部6の不揮発性メモリ62は、更に、楽曲データベースD6を記憶する。楽曲データベースD6は、楽曲の曲名と、この楽曲を実演するアーティストのアーティスト名と、この楽曲が記録されているCD、MD、及び／又はMDD53と、当該CD、MD及び／又はHDD53でこの楽曲が記録されているトラックの番号と、を互いに対応付けて格納するデータベースである。

[0034] 具体的には、楽曲データベースD6は、図3(a)に示すように、楽曲の曲名と、この楽曲を実演するアーティストのアーティスト名と、この楽曲が記録されているCD、MD及び／又はHDD53のID(識別情報)と、当該CD、MD及び／又はHDDでこの楽曲が記録されているトラックの番号と、その楽曲のテンポ、その楽曲に関する種々の属性情報(例えば、その曲の雰囲気が適した季節、時間帯、地理的位置などの情報が)が互いに対応付けて格納する。

[0035] また、楽曲データベースD6は、図3(b)に示すように、ユーザを特定するユーザ情報と楽曲の再生履歴を示す再生履歴情報を対応付けて記憶する。ユーザ情報は、ユーザ個人を特定できる情報である必要はないが、ユーザを区別できる情報である。ユーザ情報は、例えば、各ユーザの音声波形の基本ピッチの形状を示す基本ピッチ情報から構成され、音声入力部1で取得され、言語解析部2から単語データに付随して提供される。再生履歴情報は、楽曲を特定する情報と、その楽曲を再生したことを示す情報、その楽曲の再生をリピートしたことを示す情報、その楽曲の再生を途中で停止したことを示す情報、その楽曲の途中で「うるさい」などの否定的な感想を述べたことの履歴情報等を含む。これらの情報を分析することにより、ユーザの嗜好(好きな楽曲、嫌いな楽曲)を判別することができる。

[0036] エージェント処理部6の不揮発性メモリ62は、単語データベースD3を記憶する。単語データベースD3は、複数の単語のデータと、複数の単語のグルーピングを示すための単語グルーピング用のフラグとを、互いに対応付けて格納する。

[0037] 1個の単語に対応付けられているフラグは、ある概念に対応付けられてグルーピングされている。そして、フラグが所定の値(以下では、“1”)を示す場合は、このフラグに対応付けられた単語が、このフラグに対応付けられたグループに所属している。

一方、このフラグが他の値(例えば“0”)を示す場合は、この単語は、そのフラグに対応付けられたグループには所属していない。

[0038] 図4は、グルーピング用のフラグの具体例を模式的に示す図である。図4に示すように、複数の単語(図3では「再生」、「聞きたい」、「停止」、「つまらない」、「昨日」、「越天楽」、「津軽じょんがら節」、「久米利佳」)に、単語グルーピング用のフラグ(ビット)が所定数ずつ対応付けられている。また、いずれの単語に対応付けられたフラグ群についても、その最上位のフラグが「再生」という概念に対応付けられており、上位から2ビット目のフラグが「感想」という概念に対応付けられており、上位から3ビット目のフラグが「否定」という概念に、上位から4ビット目のフラグが「楽曲名」という概念に、上位から5ビット目のフラグが「アーティスト」という概念に対応付けられている。

[0039] 一方、図示するように、単語「再生」に対応付けられている上位5ビットのフラグ群の値が2進数“10000”であり、単語「聞きたい」に対応付けられている上位5ビットのフラグ群の値が2進数“11000”であり、単語「停止」に対応付けられているフラグ群の値が2進数“00100”であり、単語「つまらない」に対応付けられているフラグ群の値が2進数“01100”であり、単語「昨日」に対応付けられているフラグ群の値が2進数“00 000”であり、単語「越天楽」と「津軽じょんがら節」に対応付けられているフラグ群の値が2進数“00010”であり、単語「久米利佳」に対応付けられているフラグ群の値が2進数“00001”である。

[0040] この場合、このフラグ群は、概念「再生」の下には単語「再生」と「聞きたい」がグルーピングされており、概念「感想」の下には単語「聞きたい」及び「つまらない」がグルーピングされており、概念「否定」の下には単語「停止」及び「つまらない」がグルーピングされており、概念「楽曲名」の下には単語「越天楽」及び「津軽じょんがら節」がグルーピングされており、概念「アーティスト」の下には単語「久米利佳」がグルーピングされていることを示す。

また、単語「昨日」のように、いずれの概念にも属さないものがあつてもよい。

なお、各単語及び各概念は、例えば、処理項目データベースD4に格納されている各処理項目の「トリガ」、「判別条件」等として使用される。

[0041] エージェント処理部6の不揮発性メモリ62は、更に、処理項目データベースD4及

びワイヤデータベースD5を記憶している。

[0042] 処理項目データベースD4は、エージェント処理部6が実行する様々な処理、例えば、トリガ取得処理(TGxx)、判別処理(CNxx又はQBxx)及び入出力処理(後述するEXxx)の内容を、処理項目(ポインタ)毎に記述するデータ(処理項目データ)を格納したデータベースである。なお、「xx」は識別番号である。

[0043] 処理項目データベースD4に格納される処理項目のうち、「トリガ取得処理(TGxx)」の内容を記述するデータは、これらの処理を開始させるトリガを特定するトリガデータ(トリガとして取得するデータの内容を指定するデータ)と、後述する進行方向決定用の遷移定数k(その進行方向に遷移する程度を示すもので、後述する重み係数Jの算出基準となる定数)とを含む。

トリガデータは任意であり、例えば、窓が開いている量を示すデータや、室内の温度を示すデータや、言語解析部2より供給される上述の単語データである。トリガデータは、エージェント処理部6自身が行う処理から引き渡されるデータであってもよい。トリガデータが単語データである場合は、当該単語データが表す単語が属するグループに割り当てられている「概念」を示すデータでもよい。ただし、トリガ取得処理の内容は、例えば、複数のトリガ取得処理が互いに同一の単語を表す単語データを取得する様子がないように記述されるものとする。

[0044] 図5(a)にトリガ取得処理TGxxの例を示す。この例では、トリガTG01は、トリガとしての概念「再生」を取得する(概念「再生」にグループ化されている単語(図3の例では、単語「再生」又は「聞きたい」)を識別する)処理であり、その処理に後続する処理に進む(遷移する)か否かを決定するための遷移定数kは0.8である。図5(b)には、トリガ取得処理TG01のフローチャートを示す。

トリガ取得処理TG02は単語「MD」を取得する処理である。トリガ取得処理TG03は概念「否定」に対応するグループに属する単語を取得する(図3では、「停止」、「知らない」のいずれかを取得する)処理である。

[0045] 処理項目データベースD4に格納されている「判別処理(CNxx)」の内容を記述するデータは、判別条件と、判別結果としてとり得る結果のリストと、その判別結果を取った場合の処理の進行方向を決定するための遷移定数kと、後述する戻り方向の遷

移定数kとを、判別処理別に記述したデータを含む。

[0046] 図6(a)に判別処理CNxxの例を示す。この例では、判別処理CN01は「窓が開いているか否かを判別する処理」であり、開いていると判別したときに、それに後続する処理に進むか否かを決定するための遷移定数kが0.5、開いていないと判別したときにそれに後続する処理に進むか否かを決定するため遷移定数kが0.5である。この例のフローチャートを図6(b)に示す。図6(b)に示すノードCN01.1は、処理の開始点を示す始点ノード、ノードCN01.2は、窓が閉じていると判別したときの進行方向のノードであり、その遷移定数kは0.5である。さらに、ノードCN01.3は、窓が開いていると判別したときの進行方向のノードであり、その遷移定数kは0.5である。また、判別処理CN02は、MDプレーヤ52と、CDプレーヤ51と、HDD53のいずれが動作中かを判別する処理であり、動作中であると判別したときに後続する処理に進むか否かを決定するための遷移定数kがそれぞれ0.5である。

[0047] 「判別処理」は、判別に用いるデータを任意の取得源から取得する処理を含んでもよい。取得源としては、例えば、言語解析部2や、エージェント処理部6が実行する他の処理や、入出力対象機器群5に属する機器や、その他外部の機器・センサなどが考えられる。この場合、判別処理の内容を記述するデータは、例えば、判別に用いるデータの取得源を指定するデータを含む。

[0048] また、「判別処理」は、所定のデータを、判別に先立って所定の出力先に出力して、出力に対する応答を取得する処理を含んでもよい(この場合は、処理を示す記号をQBxxとする)。この種の判別処理としては、例えば、所定の質問を表すデータを、判別に先立って音声合成処理部3に引き渡し、その後、言語解析部2からの入力を待機する処理がある。判別処理QBxxにおいて、判別に先立って所定のデータを出力する場合、処理の内容を記述するデータは、例えば、出力するデータの内容と、このデータの出力先とを指定するデータを含む。

[0049] 図7(a)に判別処理QBxxの例を示す。この例では、例えば、判別処理QB01は、「MDを再生しますか？CDを再生しますか？HDDを再生しますか？」と利用者に問い合わせ、その応答(利用者の回答)が「MD」であったときの遷移定数kが0.5、「CD」であったときの遷移定数kが0.5、「HDD」であったときの遷移定数kが0.5であ

る。この例のフローチャートを図7(b)に示す。図7(b)に示すノードQB01. 1は、処理の開始点を示す始点ノード、ノードQB01. 2は、問い合わせに対して、「CD」を再生することが指定されたことを判別したときの進行方向のノードであり、その遷移定数kは0. 7である。また、ノードQB01. 3は、「MD」を再生することが指定されたことを判別したときの進行方向のノードであり、その遷移定数kは0. 5である。さらに、ノードQB01. 4は、「HDD」を再生することが指定されたことを判別したときの進行方向のノードであり、その遷移定数kは0. 5である。

[0050] 処理項目データベースD4に格納されている「入出力処理」の内容を記述するデータは、入力あるいは出力するデータの内容を指定するデータから構成されている。入力データ及び出力データは任意の内容を有していてよい。例えば、出力データは、音声合成処理部3を介して音声出力部4に発生させる音声の読みを表すデータや外部の機器を制御する制御信号であってもよい。また、入力データは、例えば、外部の機器から供給されるデータであってもよい。

図8(a)に入出力処理EXxxの例を示す。この例では、例えば、入出力処理EX01は、「窓を閉める」という動作であり、動作後の処理を行う進行方向の遷移定数kが0. 8である。この例のフローチャートを図8(b)に示す。図8(b)に示すノードEX01. 1は、処理の開始点を示す始点ノード、ノードEX01. 2は、処理の終了を示すノードであり、遷移定数kは0. 8である。なお、入出力処理EXxxについては、遷移定数kの設定などを行わずに、処理の終了を示すノードの選択を必須処理としてもよい。

[0051] ワイヤデータベースD5は、複数の処理(TG、CNxx、QBxx、EXxx)間の遷移を記述するデータ(以下、このデータをワイヤと呼ぶ)の集合から構成されている。ワイヤは、例えば図9に示すような書式で記述されたデータから構成されている。ワイヤWn(W1、W2...)は図示するように、先行する処理X(From (X))から後続する処理Y(To(Y))への遷移(From (X) To (Y))について、当該先行の処理(X)と、当該後続の処理(Y)と、当該遷移に対して与えられた重み係数Jと、を指定するデータである。なお、先行の処理Xが判別処理である場合は、当該判別処理のどの判別結果からの遷移であるか、まで記述される必要がある。

また、遷移元処理Xと遷移先Yは、それぞれ、各処理のノード番号で特定される。

各ワイヤの重み係数Jは、固定値ではなく、処理の進行に応じて、適宜計算され、設定されるものである。ワイヤの重み係数Jの計算は図15を参照して後述する。

[0052] エージェント処理部6は、処理項目データベースD4及びワイヤデータベースD5が全体として表しているフローを実行する。処理項目データベースD4及びワイヤデータベースD5は、例えば、図4～図9の例に基づくと、全体として図10～図13に示すような処理フローを記述することができる。

[0053] 図10に示すフローにおいては、エージェント処理部6は、トリガ処理TG01では、概念「再生」というグループに属す単語を示す単語データを言語解析部2より供給されるのを待機して、供給されるとこれを取得して判別処理CN01に引き渡す(ワイヤW1)。

エージェント処理部6は、判別処理CN01では、窓が開いているか否かを示す情報を窓開閉制御部54より取得し、開いていると判別すると入出力処理EX01に処理を移す(ワイヤW2)。入出力処理EX01では、窓開閉制御部54に、窓を閉めることを指示する制御信号を出力する。この結果、窓開閉制御部54は窓を閉める。その後、判別処理QB01に進む。なお、入出力処理EX01を行う前に、窓を閉めるか否かを問い合わせる判別処理QBxxを実行して、その応答に応じて入出力処理EX01を実行するようにしてもよい。

[0054] 即ち、この例では、エージェント処理部6は、楽曲の再生が指示されたことを認識した場合は、楽曲の再生を指示する処理に移る前に、車両の窓が開いているか否かを判別する。そして、開いていると判別した場合は、楽曲の再生の指示を行う前に、車内の遮音状況が悪いことを認識して、窓開閉制御部55に窓を閉めるよう指示するか、あるいは音声合成処理部3を介して音声出力部4に、窓を閉めるよう促すメッセージを読み上げる音声を発生させる。なお、楽曲の再生中に、「ボリュームを上げて」という意味を示す単語データが高いスコアを示す場合、ボリュームの上昇処理に先立つて窓の開閉状態を検出し、窓が開いていれば「窓が開いているよ。窓を閉める?」という音声データを音声合成処理部3で形成し、音声出力部4で出力させるようにしてもよい。

[0055] 判別処理CN01で「窓が閉じている」と判別すると、又は入出力処理EX01を終了

すると、質問を含む判別処理QB01に処理を移す(ワイヤW3、ワイヤW7)。判別処理QB01でエージェント処理部6は、まず、「MDを再生しますか？CDを再生しますか？HDDを再生しますか？」という文章を表すデータを音声合成処理部3に供給する。音声合成処理部3は、音声出力部4を介して、この文章を読み上げる音声を再生させる。

[0056] 続いて、エージェント処理部6は、言語解析部2から、単語「MD」、単語「CD」又は単語「HDD」を表すデータが供給されるのを待機する。エージェント処理部6は、該当する単語データが供給されると、このデータが単語「MD」と「CD」と「HDD」のいずれを表すかを判別する。そして、単語「MD」を表すと判別すると入出力処理EX03に処理を移し(ワイヤW5)、単語「CD」を表すと判別すると、入出力処理EX02に処理を移し(ワイヤW6)、単語「HDD」を表すと判別すると、入出力処理EX04に処理を移す(ワイヤW9)。

[0057] エージェント処理部6は、入出力処理EX02では、CD装置51に、CDの再生を開始することを指示する制御信号を出力する。エージェント処理部6は、入出力処理EX03では、MD装置52に、MDの再生を開始することを指示する制御信号を出力する。エージェント処理部6は、入出力処理EX04では、HDD装置53に、HDDの再生を開始することを指示する制御信号を出力する。

[0058] 一方、エージェント処理部6は、「MDを聞く」或いは「MDを再生」という単語を示すデータを言語解析部2より供給されるのを待機して(トリガ処理TG02)、供給されると、入出力処理EX03へと制御を移す(ワイヤW4)。

[0059] 同様に、エージェント処理部6は、「CDを聞く」或いは「CDを再生」という単語を示すデータを言語解析部2より供給されるのを待機して(トリガ処理TG03)、供給されると、入出力処理EX02に制御を移す(ワイヤW8)。

また、エージェント処理部6は、「HDDを聞く」或いは「HDDを再生」という単語を示すデータを言語解析部2より供給されるのを待機して(トリガ処理TG04)、供給されると、入出力処理EX04に制御を移す(ワイヤW8)。

[0060] 次に、図11に示すフローにおいては、エージェント処理部6は、トリガ処理TG54では、単語「停止」を示す単語データを言語解析部2より供給されるのを待機して、供給

されるとこれを取得して判別処理CN02に引き渡す(ワイヤW11)。

エージェント処理部6は、判別処理CN02では、MDプレーヤ52とCDプレーヤ51とHDDプレーヤ53のいずれが動作しているかを判別する。エージェント処理部6は、「MDプレーヤ52が動作している」と判別すると、入出力処理EX05に処理を移す(ワイヤW13)。入出力処理EX05では、MDプレーヤ52に、停止を指示する制御信号を出力する。この結果、MDプレーヤ52は動作を停止する。

[0061] エージェント処理部6は、判別処理CN02で、「CDプレーヤ51が動作している」と判別すると、入出力処理EX06に処理を移す(ワイヤW12)。エージェント処理部6は、入出力処理EX05で、CDプレーヤ51に、停止を指示する制御信号を出力する。この結果、CDプレーヤ51は動作を停止する。

エージェント処理部6は、判別処理CN02で、「HDDプレーヤ53が動作している」と判別すると、入出力処理EX07に処理を移す(ワイヤW14)。エージェント処理部6は、入出力処理EX07では、HDDプレーヤ53に、停止を指示する制御信号を出力する。この結果、HDDプレーヤ53は動作を停止する。

[0062] また、エージェント処理部6は、トリガ処理TG06で、単語「うるさい」を示す単語データを言語解析部2より供給されるのを待機して、供給されるとこれを取得して判別処理CN03に引き渡す(ワイヤW15)。

エージェント処理部6は、判別処理CN03で、MDプレーヤ52とCDプレーヤ51とHDDプレーヤ53のいずれが再生動作中であるかを判別する。「MDプレーヤ52が再生中である」と判別すると、入出力処理EX08に処理を移す(ワイヤW16)。入出力処理EX08では、MDプレーヤ52に、ボリュームを所定量下げさせる指示する制御信号を出力する。この結果、MDプレーヤ52は再生音量を低下する。

[0063] エージェント処理部6は、判別処理CN03で、「CDプレーヤ51が再生中である」と判別すると、入出力処理EX09に処理を移す(ワイヤW17)。入出力処理EX09では、CDプレーヤ51に、再生音量を低下させる制御信号を出力する。この結果、CDプレーヤ51は再生音量を低下させる。

また、エージェント処理部6は、判別処理CN03で、「HDDプレーヤ53が再生中である」と判別すると、入出力処理EX10に処理を移す(ワイヤW18)。入出力処理EX

10では、HDDプレーヤ53に、再生音量を低下させる制御信号を出力する。この結果、HDDプレーヤ53は再生音量を低下させる。

[0064] また、エージェント処理部6は、概念「曲名」又は概念「アーティスト名」を表す単語データが言語解析部2より供給されると、図12に示すトリガ取得処理TG07とTG08でこれらの単語を検出し、ワイヤW19又はW20を辿って、制御を入出力処理EX11に移す。

エージェント処理部6は、入出力処理EX11で、楽曲データベースD6を検索する。即ち、エージェント処理部6は、言語解析部2から提供された単語(楽曲名又はアーティスト名)で図3(a)に示すデータ構造の楽曲名とアーティスト名とを検索し、該当するものがあれば、該当する楽曲のMD/CD/HDD識別情報とトラック番号と抽出する。

検索処理が終了すると、エージェント制御部6は、制御を、ワイヤW21を辿って、判別処理CN04に移し、該当する楽曲が検出できたか否かを判別する。

[0065] エージェント制御部6は、該当する楽曲が検出できていれば、制御をワイヤW22を辿って、入出力処理EX12に移し、その楽曲を再生する。

[0066] 一方、該当する楽曲が検出できていなければ、制御をワイヤW23を辿って、入出力処理EX13に移し、適当な楽曲を検出できなかつた旨のメッセージを出力する。

[0067] また、ユーザが「つまんない」と発話したときに、これを検出してユーザが楽しめる音楽ソースを再生するようにすることも可能である。

この例を図13を参照して説明する。

まず、ターゲット処理TG09で、単語「つまんない」を待ち受ける。これを検出すると、ワイヤW24を辿って、判別処理CN05で、各プレーヤ51~53のステータスをチェックして、各プレーヤ51~53が再生動作を行っているか否かを判別する。いずれのプレーヤ51~53も動作を行っていない場合には、ワイヤW25を介して、問い合わせ付き判別処理QB11に移り、「オーディオをオンしますか?」というメッセージデータを音声合成部3を介して出力する。

ここで、「ハイ」、「Yes」などと一般的な返事があった場合には、ワイヤW26を介して問い合わせ付判定処理QB12に移り、「どのプレーヤをオンしますか?」等のプレー

ヤを特定させるメッセージを音声合成部3を介して出力する。

ここで、「MD」、「CD」、「HDD」などと発話されると、これが検出され、それぞれ、ワイヤW27, W28, W29を介して入出力処理EX14、EX15、EX16に移り、指示されたMD、CD、HDDを再生させる。

[0068] 一方、判別処理CN05で、いずれかのプレーヤ51、52又は53が動作を行っていると判別した場合には、ワイヤW30を介して、問い合わせ付き判別処理QB13に移り、「他のプレーヤに変更しますか？」というメッセージデータを音声合成部3を介して出力する。

ここで、単に「ハイ」、「Yes」などと一般的な返事があった場合には、ワイヤW31を介して入出力処理EX17に移り、現在再生動作を行っているプレーヤを停止して、他のプレーヤ(例えば、ランダムに選択)を起動する。

また、問い合わせ付き入出力処理QB11、QB13において、「MD」、「MDプレーヤ」、「CD」、「CDプレーヤ」、「HDD」、「HDDプレーヤ」などと具体的なプレーヤを特定する単語が発話されると、これが検出され、それぞれ、ワイヤW32, W35; W33, W36; W34, W37を介して入出力処理EX14、EX15、EX16に移り、指示されたMD、CD、HDDを再生させる。

[0069] なお、任意の処理において、再生対象の楽曲が複数検索された場合は、該当する複数の楽曲すべてを再生させてもよいし、制御対象機器群5から取得する任意のデータに基づいて、任意の基準に従い、再生する対象の楽曲を選択してもよい。

[0070] 複数の楽曲のうちから任意の曲を選択する手法を、図14(a)～図15(b)を参照して説明する。

[0071] 図14(a)の例では、エージェント処理部6は、判別処理CN04等で複数の楽曲が検索された場合に、入出力処理EX14で、車両位置特定部54から車両の走行状態に関するデータを取り込んで車両の速度を判別する。続いて、エージェント処理部6は、入出力処理EX15で、車両の走行速度に対応するテンポを有する楽曲を選択して、入出力処理EXxxで選択した楽曲を再生する。これにより、例えば、車両が高速道を走行中であるときは、テンポが速い楽曲を選択して、再生する等の動作が可能となる。

[0072] 図14(b)の例では、エージェント処理部6は、複数の楽曲が検索された場合に、入出力処理EX16で、車両位置特定部54から車両の走行状態に関するデータと地図情報を取り込んで、現在位置の地理的特徴(海の近く、山間部、田園地帯、車両が東西南北のいずれをむいているか等、)を判別する。エージェント処理部6は、入出力処理EX17で、地理的特徴に合致するタイトルや属性情報を有する楽曲をサーチして、入出力処理EXxxでこれを再生する。これにより、例えば、海沿いの道を走行中の場合に、この環境に関連付けられた属性を有する楽曲が再生される。これは、演奏者等のイメージから関連付けられてもよいし、また、曲のタイトルに「波」「海」などの言葉が含まれている楽曲を選曲するようにしてもよい。

[0073] 図15(a)の例では、エージェント処理部6は、複数の楽曲が検索された場合に、入出力処理EX18で、クロック装置56から現在の年月日時に関するデータを取り込んで、現在の時間的特徴(春夏秋冬、初夏、晚秋、…朝、昼、夕方、夜、等)を判別する。エージェント制御部6は、入出力処理EX19で、現在の時間的特徴に合致するタイトルや属性情報を有する楽曲をサーチして、入出力処理EXxxでこれを再生する。

[0074] 図15(b)の例では、エージェント処理部6は、複数の楽曲が検索された場合に、入出力処理EX20で、言語解析部2から話者のピッチ情報を取り込み、楽曲データベースD6内のユーザ情報を参照して、ユーザを特定する。

[0075] 続いて、エージェント処理部6は、入出力処理EX21で、特定した話者の履歴情報に基づいて、話者の嗜好に合致する楽曲を特定して、入出力処理EXxxでこれを再生する。

[0076] 話者の嗜好に合致する楽曲を選択する手法は任意である。例えば、過去に繰り返して再生している楽曲のジャンル、アーティスト、楽曲の年代等を判別し、これらに該当する楽曲を選択すればよい。

また、エージェント処理部6は、クロック装置56から、所定の条件に合致する日時データが供給されると、これに応答して、ネットワークを介して外部のサーバ等に接続して、楽曲データをダウンロードして、楽曲データベースD6に追加してもよい。

また、エージェント処理部6は、ユーザに告知すべき情報(該当するアーティストに

係る商品の販売スケジュールなど)を表すデータをダウンロードして、音声合成処理部3を介し、音声出力部4に、この情報を音声の形で提供させてもよい。

また、ユーザが要求したアーティストや曲名に対応する楽曲が楽曲データベースD6に保有されていない場合に、外部のサーバ等に接続し、該当するアーティストや楽曲等を楽曲データベースにダウンロードして楽曲D6に追加するようにしてもよい。

尚、ダウンロードを開始する前に「希望した楽曲が現在保有されていないので、購入してダウンロードするか？」という問い合わせのメッセージを出力するようにしてもよい。

[0077] また、エージェント処理部6は、アーティスト名を表す単語データが言語解析部2より供給されると、これに応答して、ネットワークを介して外部のショッピングサイトのサーバ等に接続し、該当するアーティストに係る商品の予約手続が可能な状態として、当該商品の予約を促すメッセージを、例えば音声合成処理部3を介し、音声出力部4に、音声の形で発させてもよい。

[0078] エージェント処理部6は、例えば言語解析部2が単語データを複数供給した場合などにおいては、複数の判別処理を並行して行う。この場合、エージェント処理部6は同一の単語を入力の対象とする処理(例えば、トリガ取得処理や、判別処理におけるデータの入力)が複数あって、該当する単語を表す単語データが言語解析部2より供給された場合は、これらの処理すべてを並行して行う。

[0079] 次に、各ワイヤの重み係数Jを計算する方法を説明する。

複数の処理がワイヤWにより連結されている場合には、注目するワイヤWの重み係数Jは、現在処理中の処理を起点として、注目しているワイヤWに至るまでの遷移の経路上の進行方向決定用の遷移定数kを順次乗算することにより求められる。

理解を容易にするため、図16(a)にフローを示す処理を想定する。

図16(a)の処理は、全体としては、単語「音楽」が入力されると、オーディオ機器が再生動作中か否かを判別し、再生中で無ければ、再生するか否かを発話者に問い合わせ、応答に応じて次の処理を実行するという処理である。

[0080] 図16(a)に示す処理では、処理TG101、CN101、QB101のいずれについても、各進行方向決定用の遷移定数kは、順方向についてはいずれも0.5、逆(戻り)方向

については、いずれも0. 1である。この場合、ワイヤは、例えば、図16(b)に示すように、定義される。

[0081] 例えば、エージェント処理部6は、処理(又は制御)がトリガ処理TG101に位置するとき(処理ポインタPPがターゲット処理TG101を指しているとき)、トリガ処理TG101を起点として、進行方向の遷移定数kを遷移経路に沿って順次乗算することにより、ワイヤW51～W55(図示していないワイヤが存在する場合にはそれらを含むすべて)のそれぞれの重み係数Jを計算し、計算結果をワイヤW51～W55に対応付けてワイヤデータベースD5に書き込む。

具体的には、トリガ処理TG101に処理が到達したとき、例えば、ワイヤW51の重み係数Jは、トリガ処理TG101の判別処理CN101に繋がるノードに割り当てられている遷移定数kの値すなわち0. 5となる。

判別処理CN101のワイヤW52の重み係数Jは、トリガ処理TG101のワイヤW51に係る遷移定数k=0. 5に判別処理CN101のワイヤW52に繋がるノードに割り当てられている遷移定数k=0. 5を乗じた結果すなわち0. 25となる。同様に、判別処理CN101のワイヤW54の重み係数Jは、トリガ処理TG101のワイヤW51に繋がるノードに割り当てられている遷移定数k=0. 5に判別処理CN101のワイヤW54に繋がるノードの遷移定数k=0. 5を乗じた結果すなわち0. 25となる。

[0082] ワイヤW53の重み係数Jは、トリガ処理TG101のワイヤW51に繋がるノードに割り当てられている遷移定数k=0. 5に、判別処理CN101のワイヤW52に繋がるノードに割り当てられている遷移定数k=0. 5を乗じた結果に更に判別処理QB101のワイヤW53に繋がるノードに割り当てられている遷移定数k=0. 5を乗じた結果、すなわち0. 125となる。同様に、ワイヤW53の重み係数Jは、トリガ処理TG101のワイヤW51に繋がるノードに割り当てられている遷移定数k=0. 5に、判別処理CN101のワイヤW52に繋がるノードに割り当てられている遷移定数k=0. 5を乗じた結果に、更に判別処理QB101のワイヤW55に繋がるノードに割り当てられている遷移定数k=0. 5を乗じた結果、すなわち0. 125となる。

[0083] このようにして、処理中の処理を基点としたときのそれぞれのワイヤの重み係数Jが計算される。よって、現在の状態が遷移すると、現在の処理を基点にしてその都度各

ワイヤWの重み係数Jが計算される。

[0084] 具体的には、現在の状態が判別処理CN101に遷移すると、ワイヤW52の重み係数Jは、ワイヤW52に繋がるノードに割り当てられている遷移定数kに等しい値0.5となり、ワイヤW54の重み係数Jは、ワイヤW52に繋がるノードに割り当てられている遷移定数kに等しい値0.5となる。また、ワイヤW53の重み係数JはワイヤW52に繋がるノードに割り当てられている遷移定数k=0.5と判別処理QB101のワイヤW55に繋がるノードに割り当てられている遷移定数k=0.5との積すなわち0.25となり、ワイヤW55の重み係数JはワイヤW52に繋がるノードに割り当てられている遷移定数k=0.5と判別処理QB101のワイヤW55に繋がるノードに割り当てられている遷移定数k=0.5との積すなわち0.25となる。さらに、逆方向(トリガ処理TG101に戻る方向に係るワイヤW51の重み係数Jは、ワイヤW51に繋がるノードに割り当てられている戻り方向のノードの遷移定数k=0.1に等しい値0.1となる。

[0085] また、現在の状態が判別処理QB101に遷移すると、ワイヤW53とW55の重み係数Jは、ワイヤW53とW55にそれぞれ繋がるノードに割り当てられている遷移定数kに等しい値0.5となる。

[0086] さらに、ワイヤW52の重み係数Jは、それに繋がる戻り方向のノードに割り当てられている戻り方向の遷移定数k=0.1がそのままワイヤW52の重み係数Jとなる。さらに、ワイヤW51の重み係数Jは、判別処理QB101のワイヤW52に繋がるノードに割り当てられている戻り方向の遷移定数k=0.1に、判別処理CN101のワイヤW51に繋がっている戻り方向のノードに割り当てられている戻り方向の遷移定数k=0.1の積の0.01となる。ワイヤW54の重み係数Jは、判別処理QB101のワイヤW52に繋がるノードに割り当てられている戻り方向の遷移定数k=0.1に、判別処理CN101のワイヤW54に繋がっているノードに割り当てられている遷移定数k=0.5の積の0.05となる。

各ワイヤWnの重み係数Jの変化の例を図15(c)に示す。

[0087] 重み係数Jの計算は、関連するフローの処理のみではなく、全てのフローの全てのワイヤについて実行され、計算された重み係数Jが各ワイヤに設定される。ここで現在の処理に関連のないワイヤについては、予め定められた低い計数値を割り当てるよう

にすればよい。しかし、特にトリガ取得処理を先行の処理とするワイヤについては、遷移定数kをある程度高く設定するようにする。こうすることによって、直前までなされていた会話と著しく異なる内容の会話にもジャンプすることが可能になる。

[0088] 次に、このように構成されたシステム全体の動作を、図17を参照して説明する。

音声入力部1と言語解析部2とは、独自に動作して、音声を取り込み、解析し、単語データをエージェント処理部6に提供する。

そして、エージェント処理部6は、判別条件に係る(1個又は複数個の)単語データが言語解析部2より供給されると、以下の処理を行う。

供給された単語を認識(識別)し(図17ステップS11)、それが、単語データベースD4に登録されている単語に相当するか否かを判別する(ステップS12)。登録されていなければ(ステップS12, No)、単語入力処理を終了する。

一方、登録されていれば(ステップS12, Yes)、その単語又はその単語の属すグループの「概念」が条件となっている処理に関し、単語の尤度Sとワイヤの重み係数Jの積を計算する(ステップS13)。

例えば、図16(a)に示すフローを実行している場合において、処理ポインタPPがトリガ処理TG101を指示しているとする。この場合の、各ワイヤの重み計数Jは図16(c-1)に示す通りである。

この状態で、スコアが80%の単語「音楽」と、スコアが50%の単語「再生」を示す単語データが入力されたと仮定する。

図16(a)と(b)に示す例では、トリガ処理TG101では、単語「音楽」が判別条件に関連し、判別処理QB101では、単語「再生」が判別条件に関連する。

図16(c-1)示すように、単語「音楽」を示す単語データを入力する処理を先行の処理とするワイヤW51の重み係数Jが0.5、単語「再生」を示す単語データを入力する処理を先行の処理とするワイヤW53の重み係数Jが0.25である。この場合、ワイヤW51及びW53について求められる尤度(スコア)Sと重み係数Jの積は、数式1及び2に示すとおりとなる。

[0089] (数1) ワイヤW51についての積J·S:「音楽」のスコアS80%×ワイヤW51の重み係数J(=0.5)=40

[0090] (数2) ワイヤW53についての積J・S:「再生」のスコアS50%×ワイヤW53の重み係数J(=0.25)=12.5

[0091] エージェント処理部6は、各単語の尤度(スコア)Sと重み係数Jとの積を求める上述の処理を、フローが有するすべてのワイヤについて行う。

続いて、エージェント制御部6は、計算された積J・Sがもっとも大きいワイヤを選択する(図17, ステップS14)。エージェント制御部6は、選択したワイヤに後続する処理に制御を進める(ステップS15)。例えば、ワイヤW51について求めた積が最も高い値を示した場合、入力された単語データは単語「音楽」を示すものであったと認識して、ワイヤW51が後続の処理としている判別処理CN101に遷移する。通常、現在の処理中の処理を起点するとワイヤの重み係数Jが比較的大きい。このため、一般的には、次の処理に移るが、従前と全く異なる単語で尤度の高いものが入力された場合には、その単語に対応する処理が開始されることもある。

[0092] エージェント処理部6は、処理が遷移すると、遷移後の状態に基づいて、各ワイヤの重み係数Jを再計算する(ステップS16)。

以後は、その処理の内容に従って処理を進める(ステップS17)。この例では、判別処理CN101を実行する。すなわち、エージェント処理部6は、窓開閉制御部55から窓の開閉の情報を取り込み、取り込んだ情報に基づいてワイヤW52とW54についての尤度Sと重み係数Jの積を求め、ワイヤW24についての尤度Sと重み係数Jの積J・Sを求めて、いずれかのワイヤWを選択して、選択したワイヤWに後続する処理を実行する。

なお、窓開閉制御部54からの窓の開閉を示す信号は、開の尤度が100%又は0%、閉の尤度が0%又は100%でもよく、また、開度に応じて変化させるようにしてもよい。そして、処理の過程で、入出力処理EXxxがあり、エージェント処理部6がその処理を実行して、音声データを出力すれば、それが音声として放音され、入出力機器群5への制御を行う処理を実行すると、その処理内容に従って、入出力機器群5への制御を行う。

[0093] トリガ取得処理TGxxからの遷移に対しては、ある程度高い遷移定数kを設定しておくとよい。具体的には、概念「再生」、単語「停止」、「うるさい」を取得するトリガ取得

処理TGxxからの遷移に対しては、例えば遷移定数k=0.9を与えておく。そうすると、該当する概念又は単語の入力を検出したときに、そのトリガ取得処理TGxxからのワイヤの積J·Sが大きくなり、これら他のワイヤについての積よりも大きくなりやすく、結果として、会話の流れとは異なる要求に対しても対応することが可能になる。

[0094] この実施の形態では、戻り方向への遷移も起こり得る。しかし、現実的には会話を戻すことは好ましくないことが多い。そこで、戻り方向の遷移定数kは、進行方向の遷移定数kに比べて低い値に設定するようにすればよい。そうすると、入力された音声から高いスコアの音声データが仮に得られても、戻り方向の遷移定数kが重み係数として書き込まれたワイヤについて求めた積は低い値となるため、戻り方向への遷移の可能性を低く抑えることができる。

また、エージェント処理部6は、求めた積の値が所定の条件に合致しないような処理(たとえば、積の値が所定値に達しないような処理)は、遷移を実行する対象から除外するよう取り扱ってもよい。

[0095] なお、例えば図9に示しているように、ワイヤは、処理項目から処理項目へという形で遷移を定義する。そして、ワイヤを図8に示すような形態で記述してワイヤデータベースD5に格納することにより、各処理項目同士の関係を、あたかもコンピュータのマクロ処理のように定義することが可能になる。これによって、各処理項目を容易に接続することができる。

[0096] また、トリガとなる処理項目は、実際には、接続されるワイヤに係る認識対象単語等(他の入力対象機器群からの入力の場合もあり得る)のスコアの判定になるので、ワイヤにおいてトリガ処理項目はワイヤの開始点として定義されず、ワイヤそのものが遷移元として定義されることになる。

[0097] 更に、上述のように各処理項目の接続関係をワイヤによって定義することによって、簡単にワイヤを追加することができる。例えば、「暑い」という音声入力の後に、ユーザが休憩することを意図して「ファミリーレストランを探して」という音声を入力する機会が多い場合、ファミリーレストランの検索処理項目に対して自動でワイヤを追加する。そうすると、ワイヤが自動で追加された後には、ファミリーレストラン検索処理項目に接続されたワイヤの遷移定数kをある程度大きくすることで、当該入力「ファミリーレストラ

ンを探して」に適切に対応することができるようになる。(ただしこの場合、エージェント処理部6は、例えばファミリーレストランの位置を示す情報を含んだ地図データ等を記憶し、あるいは外部の地図データ等にアクセスするものとする。)

このワイヤの自動追加は、ある処理項目からある処理項目へのジャンプの回数を計数し、これが所定回数に達したときに自動で行うようにすればよい。

[0098] 前述のように、エージェント処理部6は、処理項目データベースD4やワイヤデータベースD5の内容を、外部から供給される新たな処理項目データやワイヤへと更新する機能を有する。具体的には、例えば、処理項目データ及び／又はワイヤを記憶する外部のサーバが、その内容が更新されると、ネットワークを介し、エージェント処理部6に、処理項目データ及び／又はワイヤの更新がある旨を通知する。すると、エージェント処理部6はこの通知に応答して、ネットワークを介し、このサーバにアクセスして、新たな処理項目データ及び／又はワイヤをダウンロードする。そして、自己の処理項目データベースD4やワイヤデータベースD5に格納されている、古い処理項目データ及び／又はワイヤを、ダウンロードした新たな処理項目データ及び／又はワイヤへと更新する。

また、エージェント処理部6が、外部サーバにアクセスし、データベースが更新されていれば、これをダウンロードするようにしてもよい。

[0099] 外部サーバの構成例を図18に示す。

この外部サーバ100は、単語データ、処理項目データ及び／又はワイヤデータを通信ネットワークを介して複数の車両のエージェント処理部6に配信するためのものであり、制御部110、通信制御部120、単語データベース130、処理項目データベース140、ワイヤベース150、入力部160、出力部170、から構成される。

[0100] 制御部110は、例えばCPU(Central Processing Unit:中央演算処理装置)などから構成され、配信サーバ100の各部を制御するとともに、所定のプログラムを実行することで後述する各処理を実現する。

[0101] 通信制御部120は、例えば、モ뎀、ルータなどの通信装置から構成され、配信サーバ100と車両のエージェント処理部6(通信制御部)との間のネットワークを介した通信を制御する。

[0102] ネットワークの構成は任意である。例えば、専用線、公衆回線網、ケーブルテレビ(CATV)網、無線通信網、有線放送網、などを採用することができる。

[0103] 単語データベース130は、例えば、ハードディスク装置などの書換可能な記憶装置から構成され、配信サーバ100が配信する単語データをバージョン情報(例えば、タイムスタンプ)と共に蓄積する。

[0104] 処理項目データベース140は、例えば、ハードディスク装置などの書換可能な記憶装置から構成され、配信対象の処理項目データをバージョン情報(例えば、タイムスタンプ)と共に蓄積する。

[0105] ワイヤデータベース150は、例えば、ハードディスク装置などの書換可能な記憶装置から構成され、配信対象のワイヤデータをバージョン情報(例えば、タイムスタンプ)と共に蓄積している。

[0106] 配信サーバ100の管理者は、適宜、入力部160を操作して、各データベース130～150の情報を更新する。

各車両のエージェント処理部6は、通信制御部(CCU)64を介して、配信サーバ100からの更新通知を受信すると、例えば、図19(a)の処理を開始し、セッションを確立し(ステップS21)、更新されたデータをダウンロードし、自己のデータベースを更新する(ステップS22)。或いは、各車両のエージェント処理部6は、定期的或いは適宜、この配信サーバ100に通信制御部(CCU)64を介してアクセスしてセッションを確立し(ステップS31)、各データベースのバージョン情報(更新日付など)を取得する(ステップS32)。そして、両データのバージョンを比較し(ステップS33)、自己が記憶しているデータよりも新しいデータをダウンロードして、自己のデータベースにセットする(ステップS34)。

このような構成とすれば、配信サーバ100に記録されている情報を更新するだけで、配信サーバ100を利用可能なすべての車両制御システムのデータを適宜更新することができる。

なお、配信サーバ100には、データを圧縮及び暗号化して格納しておき、エージェント処理部6がダウンロードしてデータを伸張及び復号化してデータベースにセットすることが望ましい。このようにすることにより、通信の量を抑え、情報漏洩を防止できる

。

[0107] エージェント処理部6が配信サーバ100からデータをダウンロードするタイミングは上述の例に限定されず任意であり、例えば、ユーザの要求に対応する処理項目及び／又はワイヤが記憶されていない場合、これを検出してネットワークを介して配信サーバ100にアクセスし、新たな処理項目データ及び／又はワイヤをダウンロードするようにしてもよい。更には、エージェント処理部6に接続される入出力対象機器5が新たに追加された場合や、新たな機能が追加された場合に、自動でこれを検出して、上述と同様に、ネットワークを介して配信サーバ100にアクセスし、新たな処理項目データ及び／又はワイヤをダウンロードするようにしてもよい。

[0108] 尚、本発明におけるエージェント処理部6は、新たなワイヤを自動で生成する機能を有するが、ダウンロードした新たな処理項目及び既存の処理項目に対して、どのような関係のワイヤを設定するかについて、これを記述したプログラムと一緒にダウンロードするようにしてもよい。

[0109] 以上説明したこのオーディオシステムは、処理の内容を示すデータやワイヤが適切に記述されれば、制御する対象である機器や加える制御の内容を完全に特定することを必ずしも必要とせずに、ユーザが発した言語に応答し、この言語からユーザの欲求を推測し、この欲求を満たすためにどの機器にどのような制御を加えればよいかを適切に判断して、判断結果に従った制御を機器に加えることができるようになる。

[0110] また、処理項目データやワイヤは隨時新たなものへと更新されるので、エージェント処理部6の応答の仕方を変化させる余地が常にあり、ユーザはこのオーディオシステムとの対話に飽きにくい。

[0111] なお、このカーオーディオシステムの構成は上述のものに限られない。例えば、入出力対象機器群5に属する機器は、必ずしも直接にユーザの欲求を満たす結果をもたらす機器である必要はなく、例えば、外部の表示装置等を制御してユーザに特定の行動をとるよう促すメッセージを出力する機器(例えば、液晶ディスプレイ等の表示装置)からなっていてもよい。

[0112] また、単語データベースは、必ずしも単語を示すデータのみならず、複数の単語からなる語句を示すデータを単語データベースの要素として記憶するようにしてもよい

し、単語の一部あるいは音素を示すデータを単語データベースの要素として記憶するようにしてもよい。また、単語等は必ずしも特定の概念の下にグルーピングされている必要はなく、グルーピングを行う場合も、グルーピングを行うために用いられるデータは、必ずしもフラグの集合の形をとつていなくてもよい。

[0113] また、エージェント処理部6は、ワイヤに記述された遷移定数kを、過去に当該ワイヤが表す遷移を実行した数などに基づき所定の基準に従って変化させ、遷移定数係数kが変化後の値となるようにワイヤ(ワイヤの始点となるノード)に割り当てられている遷移定数kを書き換えてよい。

具体的には、例えば、ワイヤデータベースに、それぞれのワイヤについて、当該ワイヤが表す遷移が実行された回数を記憶しておく。そしてエージェント処理部6は、当該遷移が新たに行われる毎に、この回数の値を書き換えることにより、この回数の値を1ずつインクリメントし、それぞれのワイヤに記述された遷移定数kを、例えば、当該ワイヤについて記憶された回数に比例した値と書き換える。

[0114] また、エージェント処理部6は、判別処理や入出力処理において出力するデータを、これらの処理に引き渡されたデータや、これらの処理に伴って入力したデータや、その他任意の条件に従って変化させるようにしてもよい。

[0115] また、このカーオーディオシステムは、エージェント処理部6の制御に従って画像を出力するための表示装置(例えば、液晶ディスプレイ等)を備えていてもよく、エージェント処理部6は、入出力処理や判別処理において、処理毎に所定の画像を表示させるようこの表示装置を制御してもよい。

[0116] また、エージェント処理部6は、1個の入力処理や1個の判別処理において、連続して発話される等した複数の単語データを一括して取得するようにしてもよい。また、エージェント処理部6は、一括して取得した複数の単語データがどの概念の下で同一のグループ内にグルーピングされているかを特定し、特定した概念が所定の概念に合致する場合にのみ、取得した単語データの一部または全部を処理に用いるものとしてもよい。

[0117] また、エージェント処理部6は、トリガ取得処理、判別処理、入出力処理等の各種処理とワイヤとが全体として形成するフローを分担して行う、互いに接続された複数の

データ処理装置(例えば、コンピュータ等)から構成されていてもよい。この場合、解析処理部3を構成するそれぞれのデータ処理装置は、解析処理部3が実行し得るフロー全体のうち、自己が実行する可能性がある部分を表すデータを、処理項目データベースやワイヤデータベースの要素として記憶すれば十分である。そして、それぞれのデータ処理装置が記憶するデータが、当該データ処理装置が実行する部分の処理をマクロ定義するようなデータとなつていれば、複数のデータ処理装置に分散処理を行わせることも容易である。

[0118] また、このカーオーディオシステムは、音声入力部1や言語解析部2あるいは音声出力部4も複数備えていてよい。

また、音声入力部1は、たとえば、音声を表すデータが記録された記録媒体(たとえば、フロッピー(登録商標)ディスクや、CDや、MO(Magneto-Optical Disk)など)から波形信号を読み出して言語解析部2に供給する記録媒体ドライブ装置(たとえば、フロッピー(登録商標)ディスクドライブや、CD-ROMドライブや、MOドライブなど)を備えていてもよい。

[0119] 言語解析部2は、音声入力部1より供給された音声データが表す音声の話者(あるいは話者の性別など)を任意の手法で特定し、特定した話者を示す話者データをエージェント処理部3に供給するようにしてもよい。一方、エージェント処理部3は、例えば楽曲の再生を指示するたびに、楽曲データベースに、再生の対象となった楽曲が、話者データが示す話者のために再生されたことを示す履歴データを追加するようにしてもよい(あるいは、当該楽曲が当該話者のために再生された回数を示すデータを履歴データとして記憶しておき、この履歴データを1だけインクリメントするようにしてもよい)。そして、再生する楽曲を選択する場合は、このような履歴データに基づいて、楽曲を選択するようにしてもよい。

[0120] 以上、この発明の実施の形態を説明したが、この発明にかかるオーディオ機器制御装置は、専用のシステムによらず、通常のコンピュータシステムを用いて実現可能である。

例えば、入出力対象機器群5及び通信制御装置に接続されたパーソナルコンピュータに上述の音声入力部1、言語解析部2、音声合成処理部3、音声出力部4及び

エージェント処理部6の動作を実行させるためのプログラムを格納した記録媒体から該プログラムをインストールすることにより、上述の処理を実行するカーオーディオシステムを構成することができる。そして、このプログラムを実行するパーソナルコンピュータが、図1のカーオーディオシステムの動作に相当する処理として、例えば、図16に示すフローを実行するものとする。

[0121] なお、パーソナルコンピュータに上述のカーオーディオシステムの機能を行わせるプログラムは、例えば、通信回線の掲示板(BBS)にアップロードし、これを通信回線を介して配信してもよく、また、このプログラムを表す信号により搬送波を変調し、得られた変調波を伝送し、この変調波を受信した装置が変調波を復調してこのプログラムを復元するようにしてもよい。そして、このプログラムを起動し、OSの制御下に、他のアプリケーションプログラムと同様に実行することにより、上述の処理を実行することができる。

[0122] なお、OSが処理の一部を分担する場合、あるいは、OSが本願発明の1つの構成要素の一部を構成するような場合には、記録媒体には、その部分を除いたプログラムを格納してもよい。この場合も、この発明では、その記録媒体には、コンピュータが実行する各機能又はステップを実行するためのプログラムが格納されているものとする。

## 請求の範囲

[1] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、  
制御対象であるオーディオ機器の動作状態を示す動作状態情報取得手段と、  
前記音声認識手段が特定した候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、前記動作状態情報取得手段が取得した動作状態情報に基づいて、当該利用態様を実現するために当該オーディオ機器に加えるべき操作を特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、  
ことを特徴とするオーディオ機器制御装置。

[2] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、  
前記音声認識手段が特定した候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、  
ことを特徴とするオーディオ機器制御装置。

[3] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識手段と、  
前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定手段と、  
前記音声認識手段が特定した候補と、前記発話者特定手段が特定した発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を

実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御手段と、を備える、ことを特徴とするオーディオ機器制御装置。

- [4] 前記環境データは、前記オーディオ機器が使用される環境の現在位置を示すデータ又は前記オーディオ機器が使用される環境の遮音状況を示すデータから構成されている、ことを特徴とする請求項2に記載のオーディオ機器制御装置。
- [5] 前記環境データは、前記オーディオ機器が使用される環境の現在位置を示すデータ又は前記オーディオ機器が使用される環境の遮音状況を示すデータから構成されている、ことを特徴とする請求項3に記載のオーディオ機器制御装置。
- [6] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
制御対象であるオーディオ機器の動作状態を取得する動作状態取得ステップと、  
前記音声認識ステップで特定された候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、動作状態取得ステップで取得し動作状態とに基づき、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、  
ことを特徴とするオーディオ機器制御方法。
- [7] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
前記音声認識ステップで特定された候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、  
ことを特徴とするオーディオ機器制御方法。
- [8] 音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当

該音声が表す語句の候補を特定する音声認識ステップと、

前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定ステップと、

前記音声認識ステップで特定された候補と、前記発話者特定ステップで特定された発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとにに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、より構成される、

ことを特徴とするオーディオ機器制御方法。

[9] コンピュータに、

音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、

制御対象であるオーディオ機器の動作状態を取得する動作状態取得ステップと、

前記音声認識ステップで特定された候補に基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、動作状態取得ステップで取得し動作状態とにに基づき、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、を実行させる ことを特徴とするコンピュータプログラム。

[10] コンピュータに、

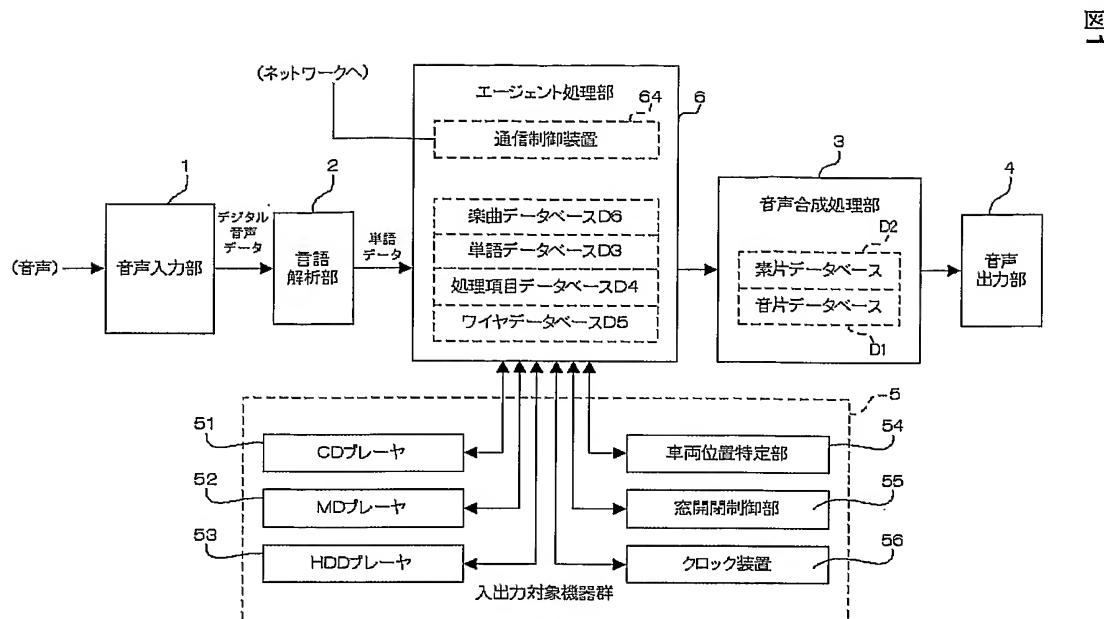
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、

前記音声認識ステップで特定された候補と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとにに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器

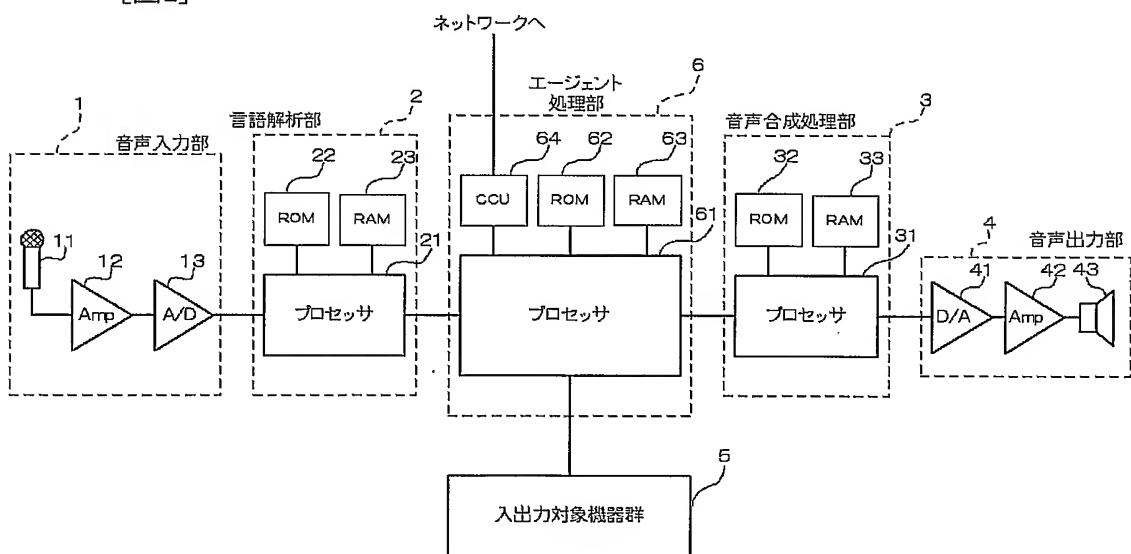
制御ステップと、  
を実行させることを特徴とするコンピュータプログラム。

[11] コンピュータに、  
音声を表す音声データを取得し、当該音声データに音声認識を施すことにより、当該音声が表す語句の候補を特定する音声認識ステップと、  
前記音声データに基づき、前記音声の発話者又は発話者の属性を特定する発話者特定ステップと、  
前記音声認識ステップ特定された候補と、前記発話者特定ステップで特定された発話者又は発話者の属性と、外部のオーディオ機器が使用される環境の状況を示す環境データ及び／又は当該オーディオ機器の動作状態を示す動作状態データとに基づき、前記音声の発話者が欲する当該オーディオ機器の利用態様と、当該利用態様を実現するために当該オーディオ機器に加えるべき操作とを特定し、特定した操作を加えるため前記オーディオ機器を制御するオーディオ機器制御ステップと、  
を実行させるコンピュータプログラム。

[図1]



[図2]



[図3]

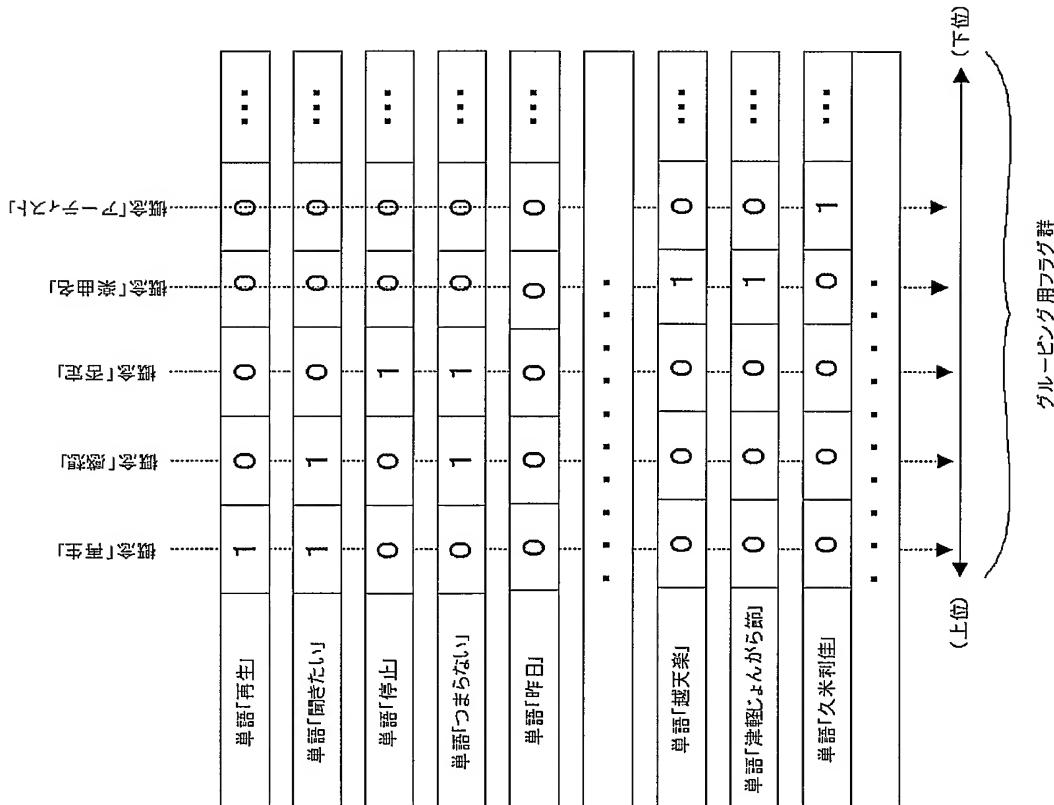
(a)

楽曲名	アーティスト	MD／CD識別情報	トラックNo	テンポ	属性情報
-----	--------	-----------	--------	-----	------

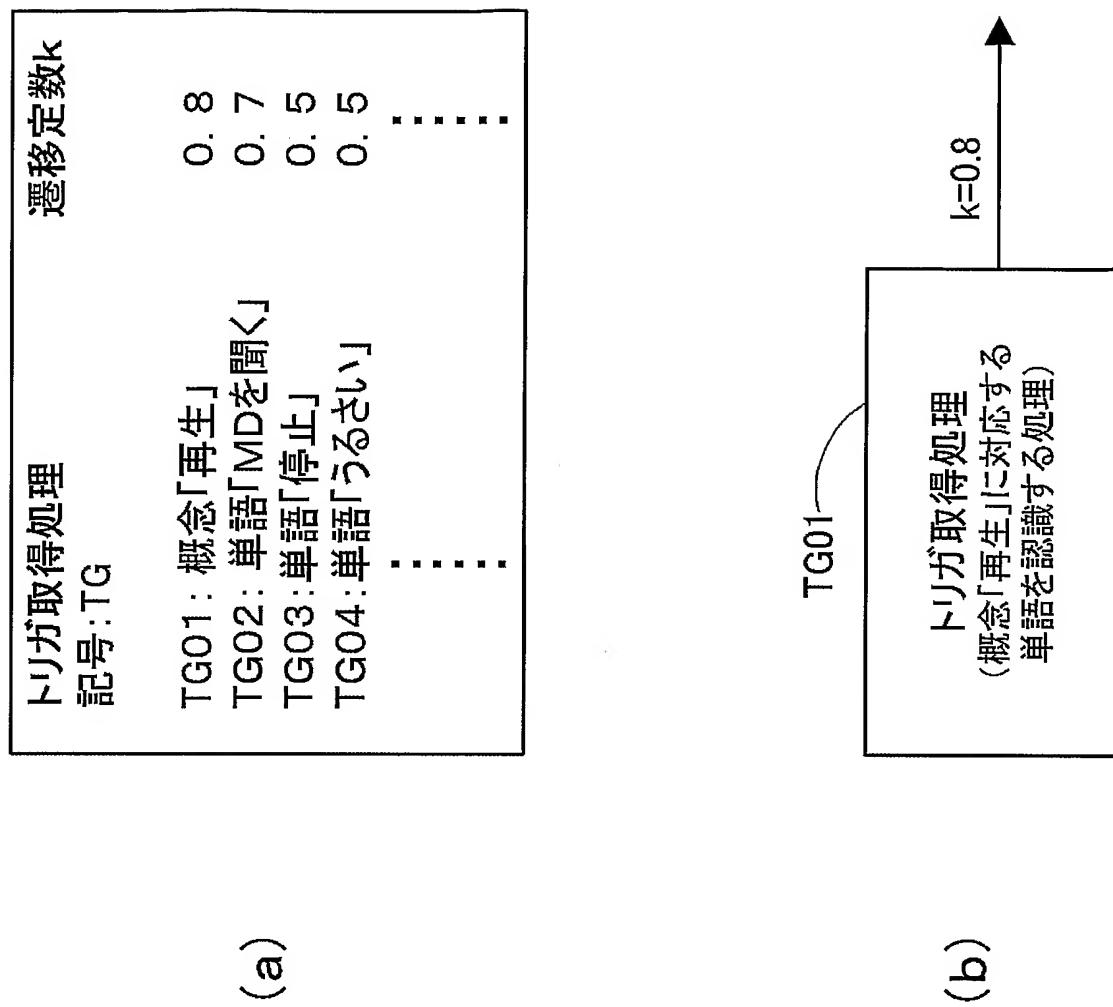
(b)

ピッチA	再生履歴データ (再生データ 停止データ、 うるさいデータ)
ピッチB	履歴データ

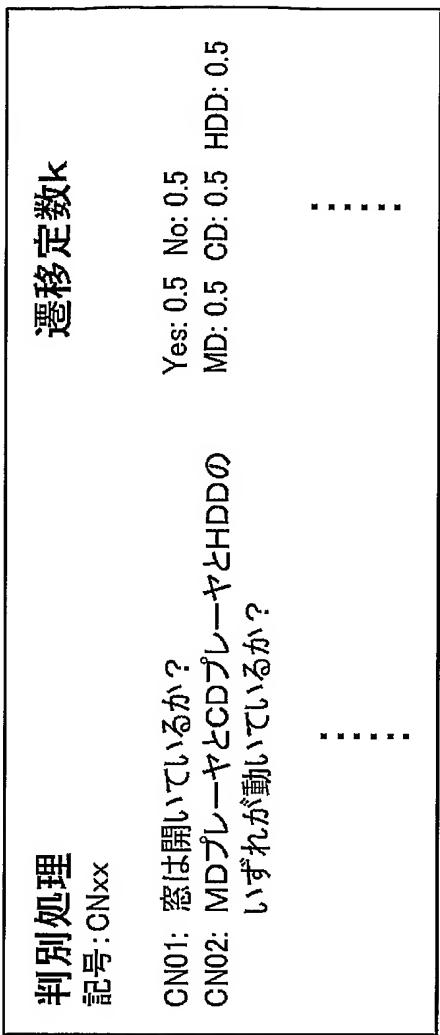
[図4]



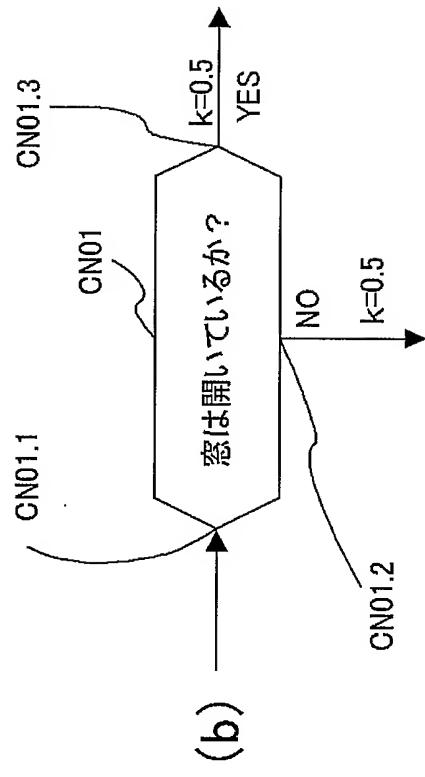
[図5]



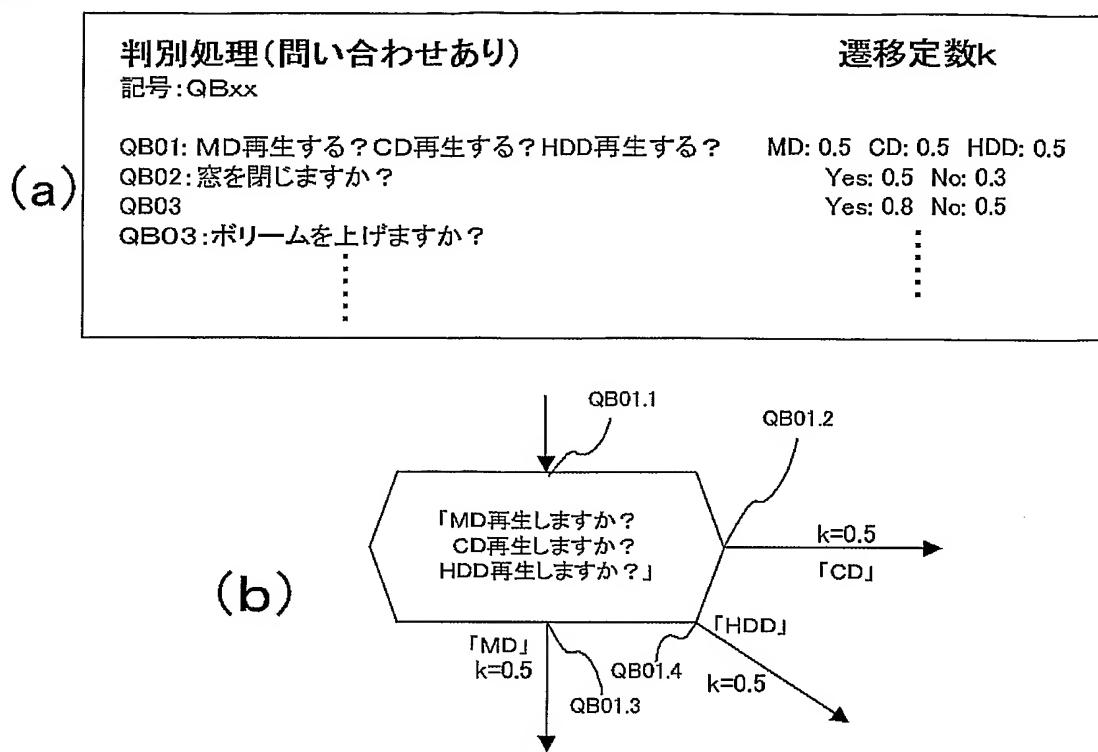
[図6]



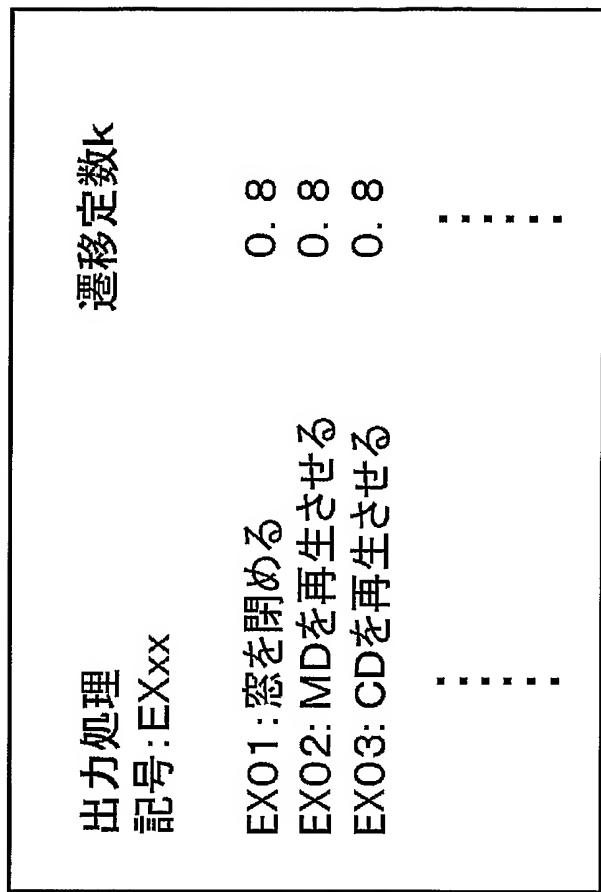
(a)



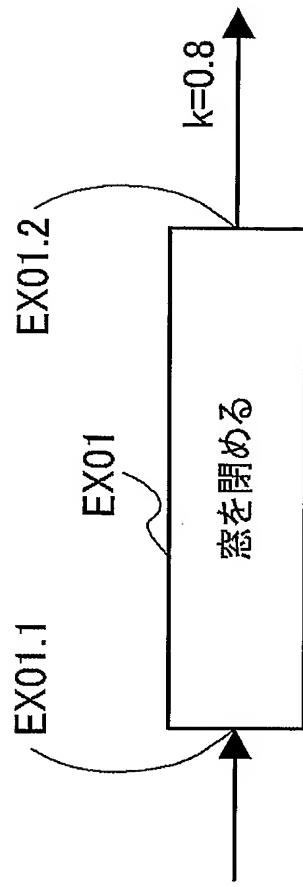
[図7]



[図8]



(a)

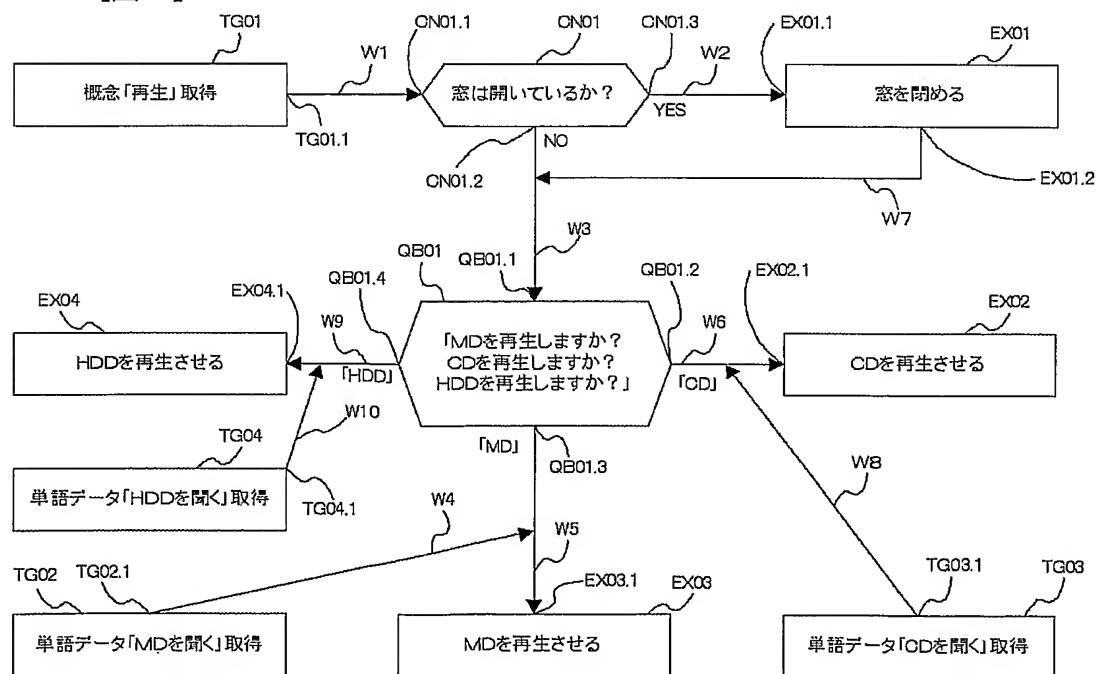


(b)

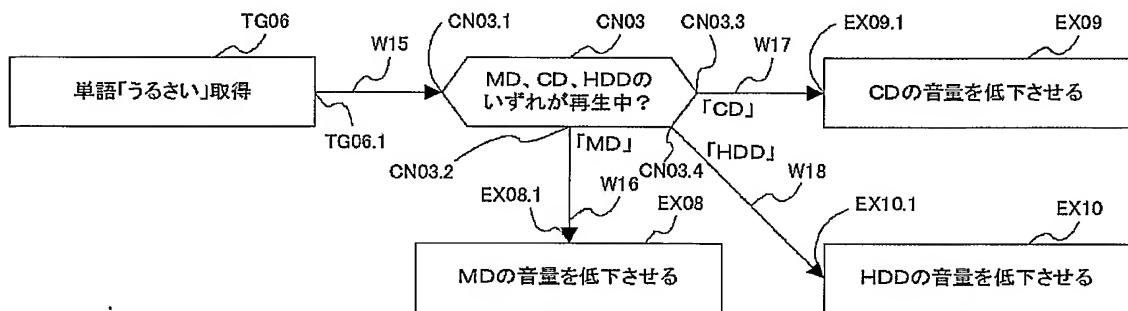
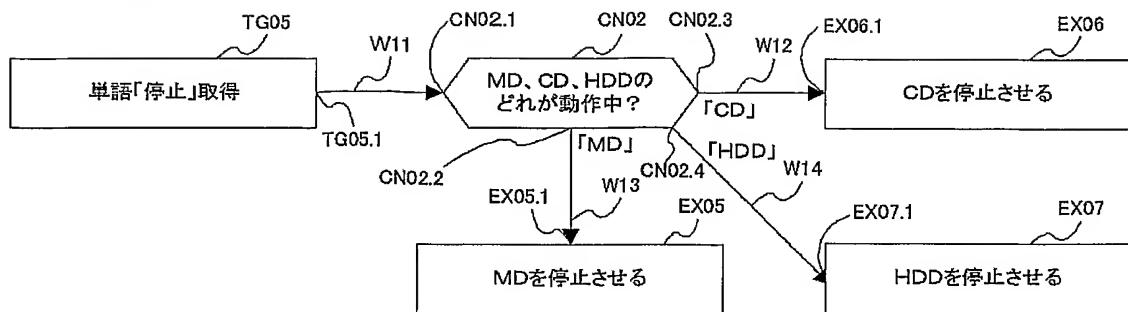
[図9]

ワイヤ Wn	先行 ノード	後行 ノード	重み 係数J
W1 = From (TG01.1) To (CN01.1), 1.0			
W2 = From (CN01.3) To (EX01.1), 0.5			
W3 = From (CN01.2) To (QB01.1), 0.5			
W4 = From (TG02.1) To (EX03.1), 1.0			
W5 = From (QB01.3) To (EX03.1), 0.25			
W6 = From (QB01.2) To (EX02.1), 0.25			
W7 = From (EX01.2) To (QB01.1), 0.5			
W8 = From (TG03.1) To (EX02.1), 1.0			
W9 = From (QB01.4) To (EX04.1), 0.25			
W10 = From (TG04.1) To (CN04.1), 1.0			
W11 = From (TG05.1) To (CN02.1), 1.0			
W12 = From (CN02.3) To (EX06.1), 0.5			
W13 = From (CN02.2) To (EX05.1), 0.5			
W14 = From (CN02.4) To (EX07.1), 0.5			
W15 = From (TG06.1) To (CN03.1), 1.0			
W16 = From (CN03.2) To (EX08.1), 0.5			
W17 = From (CN03.3) To (EX09.1), 0.5			
W18 = From (CN03.4) To (EX10.1), 0.5			
W19 = From (TG06.1) To (EX11.1), 1.0			
W20 = From (TG07.1) To (EX11.1), 1.0			
W21 = From (EX11.2) To (CN04.1), 1.0			
W22 = From (CN03.2) To (EX12.1), 0.5			
W23 = From (CN03.3) To (EX13.1), 0.5			
W24 = From (TG09.1) To (CN05.1), 1.0			
W25 = From (CN05.3) To (QB11.1), 1.0			
W26 = From (QB11.2) To (QB12.1), 0.5			
W27 = From (QB12.2) To (EX14.1), 0.25			
W28 = From (QB12.3) To (EX15.1), 0.25			
W29 = From (QB12.4) To (EX16.1), 0.25			
W30 = From (CN05.2) To (QB13.1), 1.0			
W31 = From (QB13.2) To (EX17.1), 0.5			
W32 = From (QB11.3) To (EX14.1), 0.5			
W33 = From (QB11.4) To (EX15.1), 0.5			
W34 = From (QB11.5) To (EX16.1), 0.5			
W35 = From (QB13.3) To (EX14.1), 0.5			
W36 = From (QB13.4) To (EX15.1), 0.5			
W37 = From (QB13.5) To (EX16.1), 0.5			

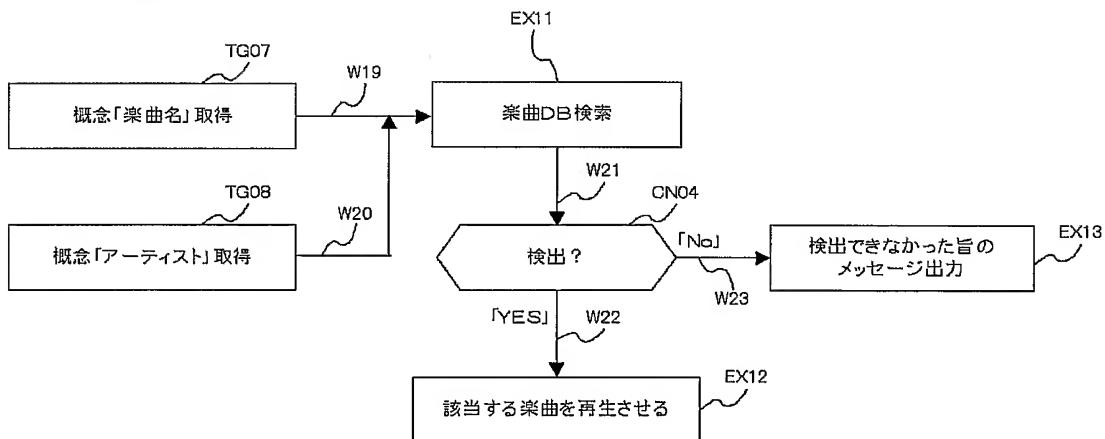
[図10]



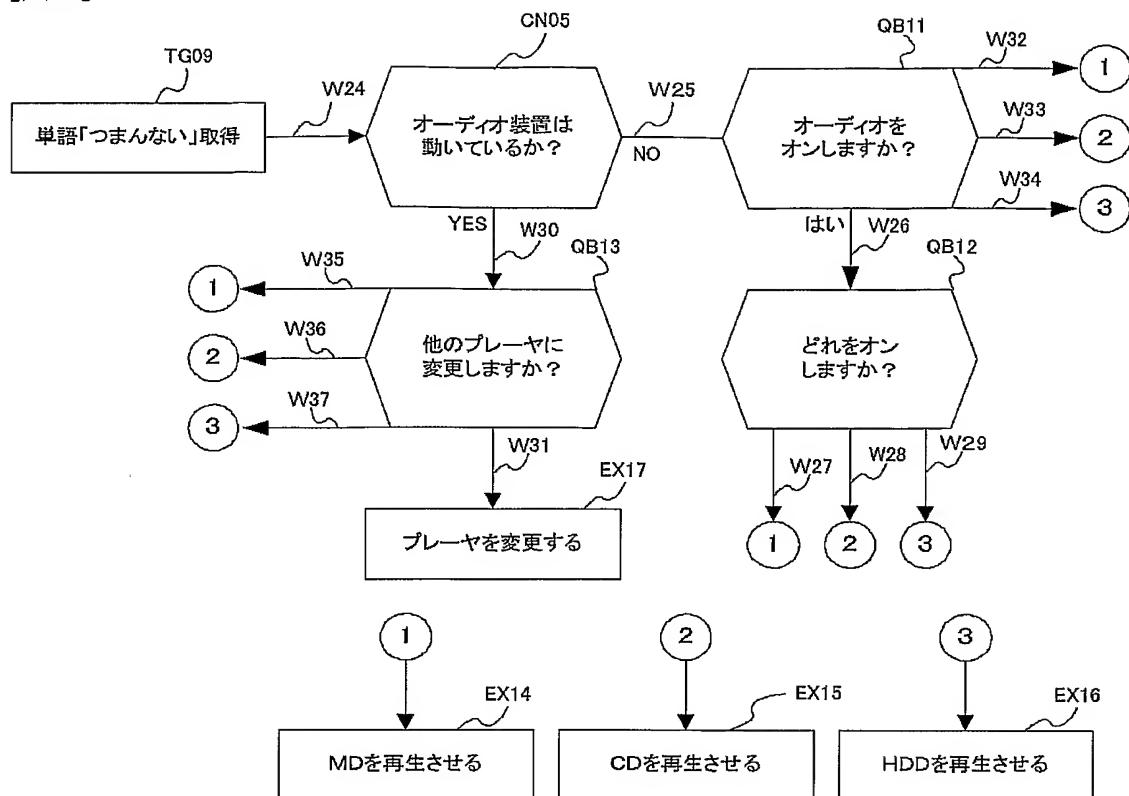
[図11]



[図12]



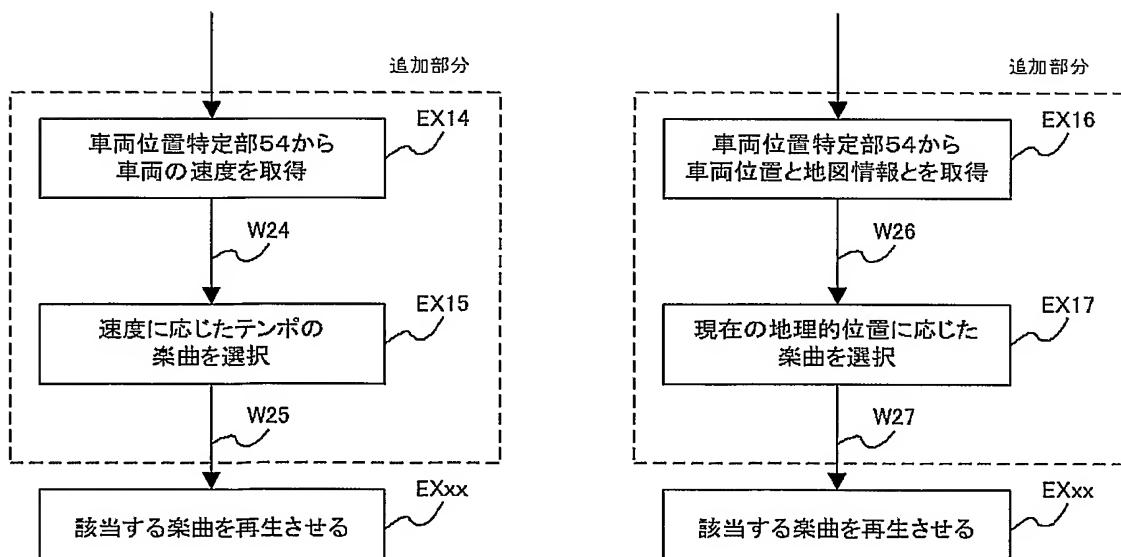
[図13]



[図14]

楽曲を選択する処理

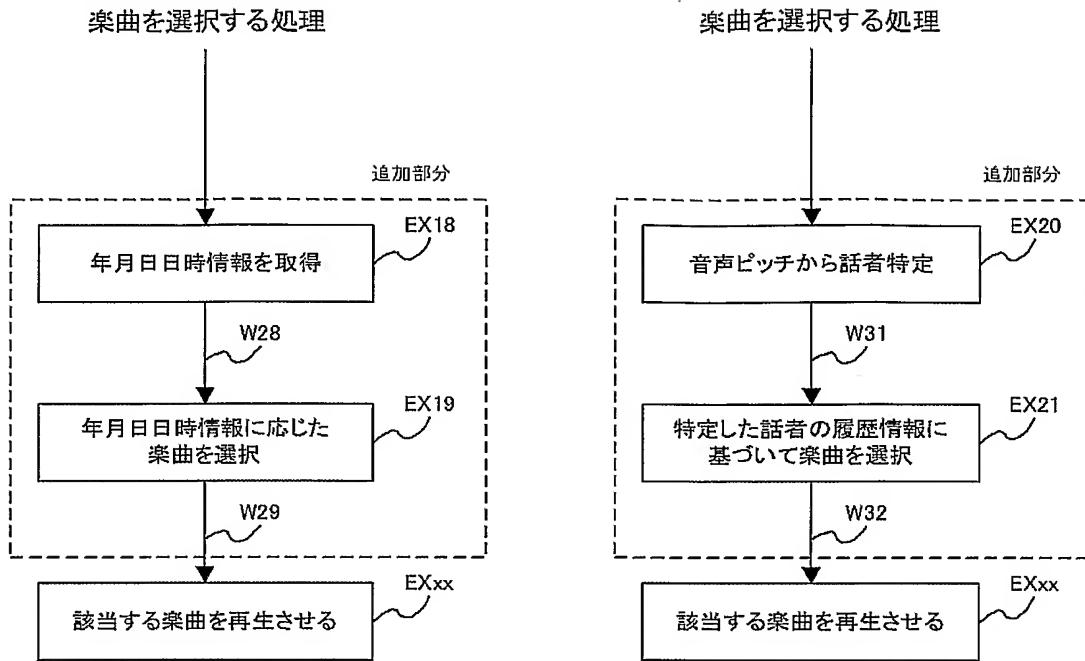
楽曲を選択する処理



(a)

(b)

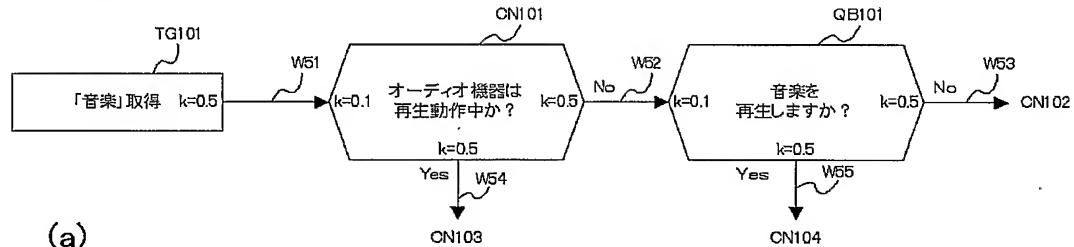
[図15]



(a)

(b)

[図16]



(a)

W51 = From (TG101. 1) To (CN101.1)  
 W52 = From (CN101.3) To (QB101.1)  
 W53 = From (QB101.3) To (CN102.1)  
 W54 = From (CN101.2) To (CN103.1)  
 W55 = From (QB101.2) To (CN104.1)

(b)

処理がTG101にあるとき

W51 = 0.5  
 W52 = 0.25  
 W53 = 0.125  
 W54 = 0.25  
 W55 = 0.125

(b-1)

処理がCN101にあるとき

W51 = 0.1  
 W52 = 0.5  
 W53 = 0.25  
 W54 = 0.5  
 W55 = 0.25

(b-2)

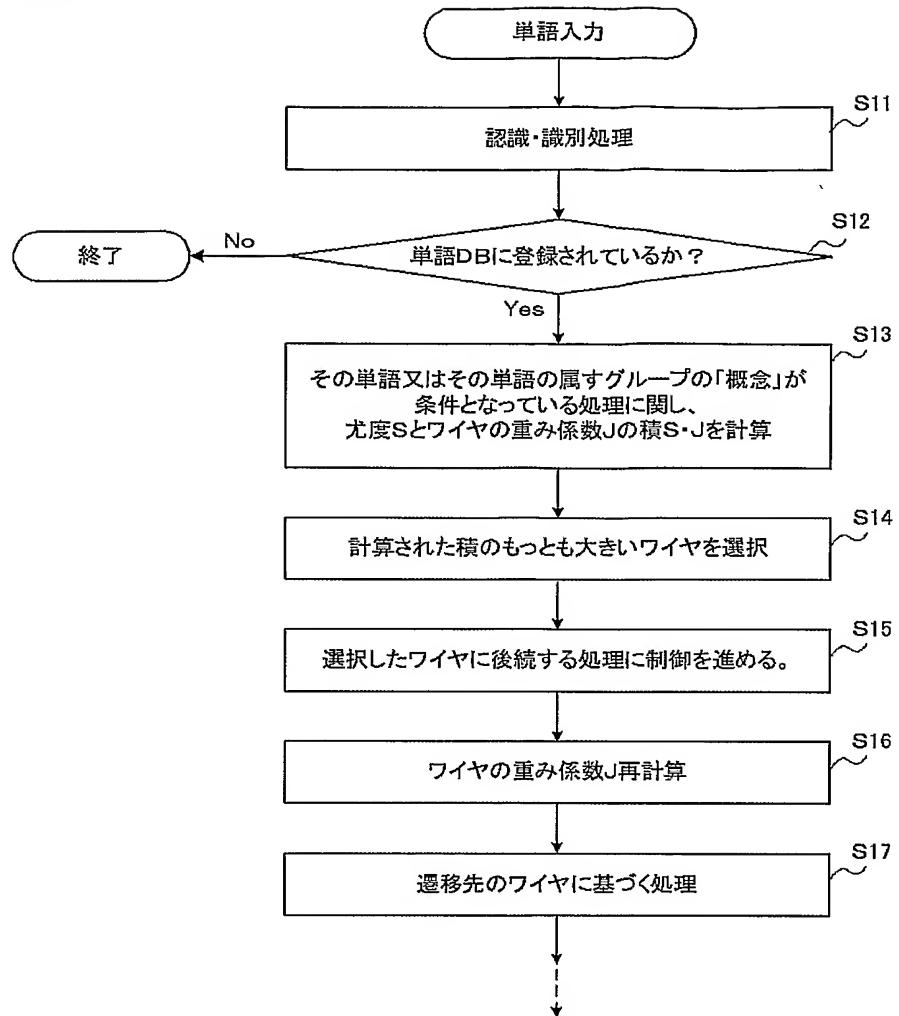
処理がQB101にあるとき

W51 = 0.01  
 W52 = 0.1  
 W53 = 0.5  
 W54 = 0.05  
 W55 = 0.5

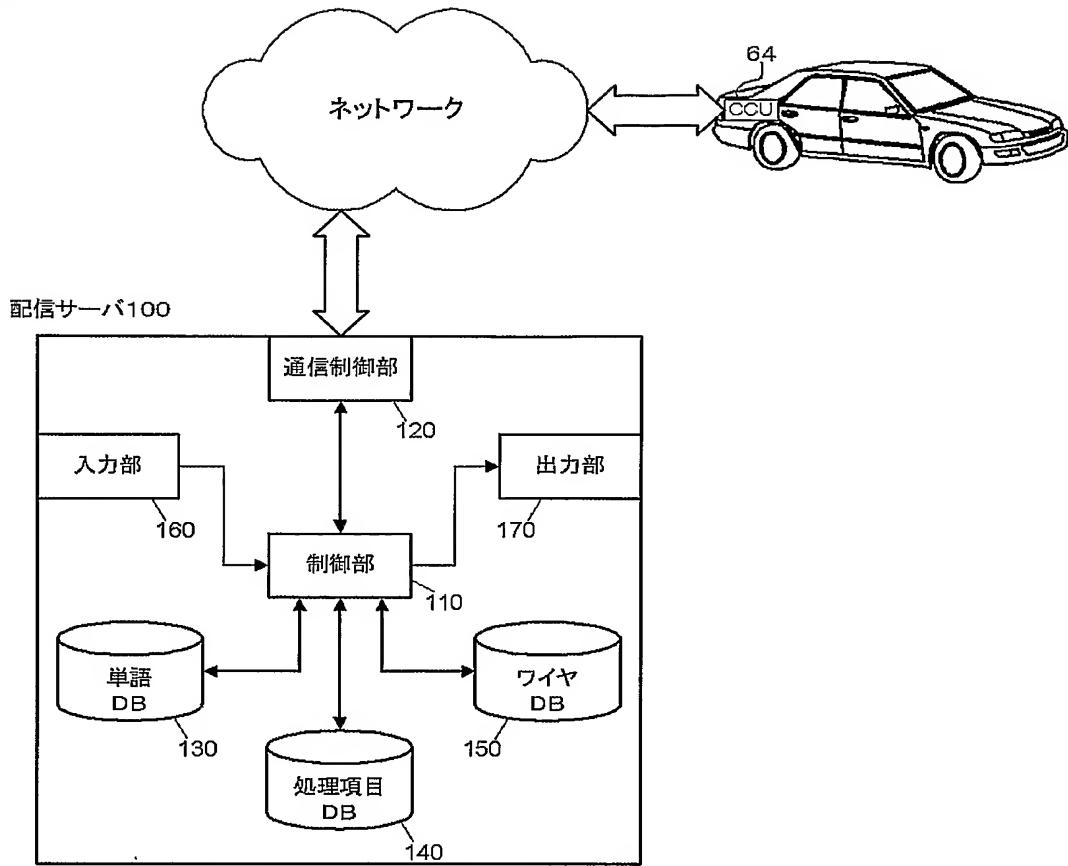
(b-3)

(c)

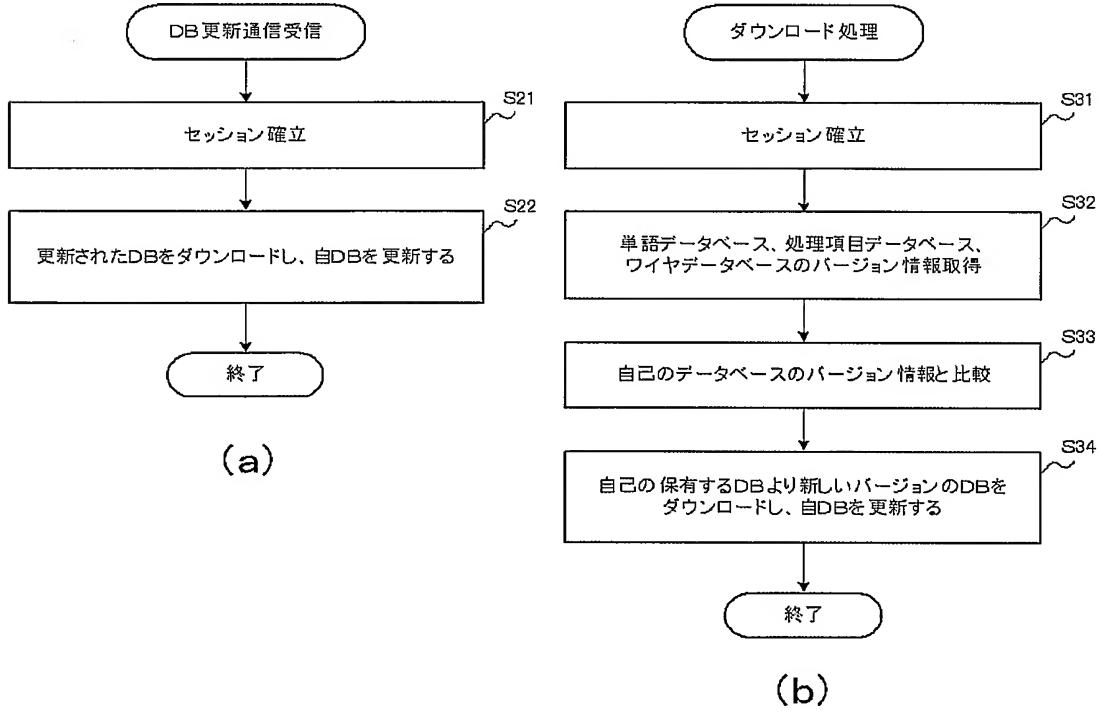
[図17]



[図18]



[図19]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2004/004007

## A. CLASSIFICATION OF SUBJECT MATTER

Int.C1<sup>7</sup> G10L15/00, 15/22, 15/28, G11B20/10

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.C1<sup>7</sup> G10L15/00, 15/22, 15/28, G11B20/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2000-181500 A (Equos Research Co., Ltd.), 30 June, 2000 (30.06.00), (Family: none)	1, 6, 9
Y	JP 2002-165298 A (Matsushita Electric Industrial Co., Ltd.), 07 June, 2002 (07.06.02), (Family: none)	2-5, 7, 8, 10, 11
Y	JP 2003-140664 A (Fujitsu Ten Ltd.), 16 May, 2003 (16.05.03), (Family: none)	2, 7, 10
		3-5, 8, 11

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
18 April, 2005 (18.04.05)Date of mailing of the international search report  
10 May, 2005 (10.05.05)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> G10L15/00, 15/22, 15/28, G11B20/10

## B. 調査を行った分野

## 調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> G10L15/00, 15/22, 15/28, G11B20/10

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	J P 2000-181500 A (株式会社エクオス・リサーチ) 2000.06.30 (ファミリーなし)	1、6、9
Y	J P 2002-165298 A (松下電器産業株式会社) 2002.06.07 (ファミリーなし)	2~5、7、8、 10、11
Y	J P 2003-140664 A (富士通テン株式会社) 2003.05.16 (ファミリーなし)	2、7、10
		3~5、8、 11

\* C 欄の続きにも文献が列挙されている。

\* パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

- 「A」特に関連のある文献ではなく、一般的技術水準を示すもの
- 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
- 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
- 「O」口頭による開示、使用、展示等に言及する文献
- 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

- 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
- 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
- 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
- 「&」同一パテントファミリー文献

## 国際調査を完了した日

18.04.2005

## 国際調査報告の発送日

10.5.2005

## 国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

桂原 雄一

5C 3148

電話番号 03-3581-1101 内線 3541